4-30-2018

Differences in Cognitive Flexibility Within the Primate Lineage and Across Human Cultures: When Learned Strategies Block Better Alternatives

Sarah M. Pope

Recommended Citation

Pope, Sarah M., "Differences in Cognitive Flexibility Within the Primate Lineage and Across Human Cultures: When Learned Strategies Block Better Alternatives." Dissertation, Georgia State University, 2018.
https://scholarworks.gsu.edu/neurosci_diss/33

DIFFERENCES IN COGNITIVE FLEXIBILITY WITHIN THE PRIMATE LINEAGE AND ACROSS HUMAN CULTURES:
WHEN LEARNED STRATEGIES BLOCK BETTER ALTERNATIVES

by

SARAH MICHELLE POPE

Under the Direction of William D. Hopkins, PhD

ABSTRACT

By applying learned rules, humans are able to accurately solve many problems with minimal cognitive effort; yet, this sort of habit-based problem solving may readily foster a type of cognitive inflexibility termed 'cognitive set'. Cognitive set occurs when an alternative – even more efficient – strategy is masked by a known, familiar solution. In this research, I explored how cognitive set differs between primate species and across human cultures, using a nonverbal computerized 'LS-DS' task, which measures subjects' ability to depart from a three-step, learned strategy (LS) in order to adopt a more efficient, one-step, direct strategy (DS or 'the shortcut'). First, I compared baboons', chimpanzees', and humans' abilities to break cognitive set and found that all baboon and chimpanzee subjects used the DS shortcut when it became available; yet, humans exhibited a remarkable preference for the LS. Next, in an effort to elucidate *how* cognitive set occludes alternative strategies, I tracked human participants' eye movements to identify whether better solutions are a) visually overlooked or b) seen but disregarded. Although human subjects saw the shortcut, they did not use it until their conceptualization of the problem constraints were altered. Lastly, to further distinguish between perceptual and conceptual influences on cognitive set, I compared shortcut-use between Westerners and the semi-nomadic Himba of northern Namibia. This study found that susceptibility to cognitive set varied across human cultures and presented further evidence that problem conceptualization, and not perceptual processing, influences individuals' ability to break set and use the alternative. Overall, this research provides a novel comparison of cognitive flexibility within the primate lineage and across human cultures. The implications for set-promoting influences, including the potentially mechanizing problem-solving methods typical of Western education, are discussed.

INDEX WORDS: Cognitive flexibility, Evolution, Cross-cultural, Perception, Primates

DIFFERENCES IN COGNITIVE FLEXIBILITY WITHIN THE PRIMATE LINEAGE AND ACROSS HUMAN CULTURES:

WHEN LEARNED STRATEGIES BLOCK BETTER ALTERNATIVES

by

SARAH MICHELLE POPE

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

2018

DIFFERENCES IN COGNITIVE FLEXIBILITY WITHIN THE PRIMATE LINEAGE AND ACROSS HUMAN CULTURES:

WHEN LEARNED STRATEGIES BLOCK BETTER ALTERNATIVES

by

SARAH MICHELLE POPE

Committee Chair: William Hopkins

Committee: Joël Fagot

Adrien Meguerditchian

David Washburn

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
May 2018

DEDICATION

To my mother, Sandra W. Doncaster,
who most fervently supported my *monkey scientist* dreams.

Thank you for teaching me,
for letting me learn the hard way sometimes,
for holding my hand,
and for challenging me to explore.

You are immeasurably missed.

ACKNOWLEDGEMENTS

Many people have contributed to the synthesis of this research. I cannot hope to adequately thank them all on a single page, but here goes.

To Bill Hopkins, who graciously provided me with the rare opportunity and resources to research what interested me most. For that, I am eternally grateful. Thank you for all of your encouragement, critiques, advice, and support.

To Joël Fagot, who welcomed me into his lab and truly shaped my research path. Thank you for lending your knowledge, time, and resources to my research pursuits. The learning environment, which you provided, was truly inspirational.

To Adrien Meguerditchian, who made this research and my cotutelle possible. Thank you for always having an idea to discuss with me. Thank you for challenging me to think critically and for your ever-enthusiastic feedback, it provided encouragement when I needed it most. Also, thank you for not being terribly angry when I got a speeding ticket in Switzerland with your car.

Additionally, I want to thank Lindsay Mahovetz for all the time she spent helping to debug the program and generally for being the kindest person I know. To Michelle Autrey, thank you for the countless hours collecting data / avoiding projectile chimp poo with me. To each of the inspiring people who have contributed to this project, including: David Washburn, Sarah Davis, Arianne Ribeiro, Nadia deKoekkoek, thank you.

Lastly, to my family, but especially Michael Pope, Jesse and Philip Coffey, and Philip Doncaster, your enormous contributions over the years are what have made this possible. I will never be able to fully express my gratitude. Thanks for putting up with me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CHAPTER 1

S-R……………………………………………………………Stimulus-Response
RT……………………………………………………………Response Time
WCST………………………………………………Wisconsin Card Sorting Task
DCCS……………………………………………Dimensional Change Card Sort
CSST………………………………………………Conceptual Set Shifting Task
WEIRD……………………Western, Educated, Industrialized, Rich, Democratic
LS-DS………………………………………Learned Strategy – Direct Strategy
Square1……1st Square to flash red in the demo, and the correct LS 1st response.
Square2….2nd Square to flash red in the demo, and the correct LS 2nd response.
Triangle…Correct 3rd LS response or 1st response when using the DS or shortcut.
LS………………………………Learned Strategy, Sqaure1→ Square2→Triangle
BASE…………..…………Baseline Condition, in which the Triangle is hidden.
PROBE……Probe Condition, in which the Triangle is visible throughout the trial.
DS……………………………………Direct Strategy, 1st response is Triangle
SS………………………………………Switch Strategy, Square1 → Triangle

CHAPTER 2

ALDM…………………………………Automated Learning Device for Monkeys
DSer……………………………………………………………DS Strategy-User
LSer……………………………………………………………LS Strategy-User

CHAPTER 3

RT1…………………………………………Response Time for the 1st Response
BASE *stay*………BASE trial in which subjects' repeated their previous strategy.
BASE *switch*…………………BASE trial in which subjects' switched strategies.
PROBE *stay*……PROBE trial in which subjects' repeated their previous strategy.
PROBE *switch*………………PROBE trial in which subjects' switched strategies.
BASE *same*…………………………………BASE trial preceded by a BASE trial.
BASE *different*……………………………BASE trial preceded by a PROBE trial.
PROBE *same*……………………………PROBE trial preceded by a PROBE trial.
PROBE *different*…………………………PROBE trial preceded by a BASE trial.

CHAPTER 4

Informed…………………Subjects who watched the video demonstrating the DS
Control………………… Subjects who watched the video demonstrating the LS
PRE…………The first block of 48 trials, occurring prior to watching the video.
POST………The second block of 48 trials, occurring after watching the video.
Look LS………………………Subjects used the LS after fixating on the Triangle
No Look LS…………… Subjects did not use the LS after fixating on the Triangle

CHAPTER 5

PRE………The first block of 48 trials, occurring prior to receiving the prompt.
POST………The second block of 48 trials, occurring after receiving the prompt.

CHAPTER 6

PFC……………………………………………………………Prefrontal Cortex
VLPFC…………………………………………Ventrolateral prefrontal Cortex

1 INTRODUCTION TO COGNITIVE FLEXIBILITY IN PRIMATES AND ACROSS HUMAN CULTURES

1.1 Introduction to Cognitive Flexibility

1.1.1 Adaptive Significance of Flexible Behavior

Flexible strategy-use is the foundation of adaptive problem solving. Situations requiring decisive action emerge almost continuously when navigating dynamic environments and an individual's response to such shifting social and ecological inputs can be profoundly consequential. Fitness, and even survival, often hinge upon adaptive problem solving in which a response is not only determined by variable stimuli but also by shifting contexts (Potts, 2012; Reader, Morand-Ferron, & Flynn, 2016; Tebbich, Griffin, Peschl, & Sterelny, 2016). Accordingly, cognitive flexibility is considered a nontrivial component of intelligent behavior (Buttelmann & Karbach, 2017; Genovesio & Wise, 2008; Ionescu, 2012; Rosati, 2017; Stoet & Snyder, 2008).

Here, I will broadly define cognitive flexibility as the ability to incorporate both known solutions and innovated or acquired novel solutions in a contextually appropriate manner (modified from Buttelmann & Karbach, 2017; and Lehner, Burkart, & Schaik, 2011). Thus, flexible responses must integrate external environmental cues with internal inputs, such as past experience, and in the case that a previous strategy is no longer the most appropriate, flexible behavior requires inhibiting that previous response and switching to a more efficient strategy.

During development, humans acquire and refine a suite of cognitive skills that eventually enable successful, and self-sufficient, navigation of their environment. This suite of skills, termed executive functioning, or the processes underlying the conscious control of action

(Zelazo, 2008), are commonly identified as being comprised of three components: inhibition, working memory, and switching (Miyake & Friedman, 2012). From early childhood to adolescence, executive functioning develops rapidly, which has been associated with developmental changes in the prefrontal cortex (reviewed in Buttelmann & Karbach, 2017; Genovesio & Wise, 2008; Thompson-Schill, Ramscar, & Chrysikou, 2009). Accordingly, children, adolescents and adults exhibit differing degrees of cognitive flexibility across a range of tasks; yet both direct (increase in age = increase in flexibility; Cunningham, 1965; Zelazo, 2008) and inverse (increase in age = decrease in flexibility; Ardiale & Lemaire, 2012; Defeyter & German, 2003; German & Defeyter, 2000; Gopnik, Griffiths, & Lucas, 2015; Lemaire & Leclere, 2014; Lucas, Bridgers, Griffiths, & Gopnik, 2014; Luchins, 1942) relationships between cognitive flexibility and age are observed (however, see: Chelune & Baer, 1986; Diamond & Kirkham, 2005). Furthermore, cognitive flexibility has been linked to academic achievement (Buttelmann & Karbach, 2017; Cole, Duncan, & Blaye, 2014; Titz & Karbach, 2014), social skill (Buttelmann & Karbach, 2017), health (Besnard & Cacitti, 2005; Masley, Roetzheim, & Gualtieri, 2009), and is found to differ between some clinical populations such as in alcoholism (Trick, Kempton, Williams, & Duka, 2014) and Autism Spectrum Disorder (Yeung, 2015). Still, flexible behavior is not unequivocally beneficial.

The interplay between flexible strategy updating and maintenance of extant solution strategies presents an interesting opposition. On one hand, if an existing representation is too concrete and impervious to contextual inputs, then newly relevant information is ignored, resulting in a familiar but potentially inefficient response. On the other hand, if the representation is too easily perforated, then response efficiency is impaired by distractibility (Cools, 2008; Hommel & Colzato, 2017; Roberts, 2008). For example, if a foraging animal is unable to adapt flexibly to shifting environments, such as a new stream that must be crossed in order to reach food, they may starve. Yet, if behavior is too flexible, the same animal may attempt to cross an insurmountable obstacle, with the same fatal result. Thus, optimal behavior requires a balance between exploiting a known solution and exploring alternatives (Brosnan & Hopper, 2014; Chrysikou et al., 2013).

1.1.2 Abstract Problem Solving

Problems range wildly in complexity, as do possible solutions. In simple cases, an appropriate response may be acquired through conditional associations or trial and error learning (Petrides, 2008). This occurs when a stimulus and an action produce a reward with some consistency, thus favoring that specific response in the presence of that specific stimulus. These Stimulus-Response (S-R) associations often require an extended learning period but, once developed, are characterized by faster, more consistent, responses and reduced distractibility (Smith & Graybiel, 2016). However, such automaticity, although certainly beneficial in some situations (red traffic light → stop), cannot be generalized to instances in which that specific stimulus is not present (pedestrians crossing the road → stop). Thus, countless S-R associations would be required to fully understand a problem space, placing an impractical load on one’s cognitive processing.

Abstract response contingencies overcome this limitation by providing a response framework that can be generalized across situations (Miller & Buschman, 2008). Derived from the Latin word abtrahere, which translates to “to drag away,” abstractions are not fixed to a specific stimulus (Christoff & Keramatian, 2008). Thus, whereas S-R associations reduce cognitive load by automating behavior, abstract rules reduce cognitive load by applying a single solution to multiple problems. Yet, selecting an appropriate response strategy, in and of itself, leaves room for error. Simply put, an action is only as effective as the strategy that guides it.

> *"Decisions based on abstractions require a number of coordinated processes, including the top-down biasing of inputs to the prefrontal cortex; the categorization of contextual information, including sensory inputs, memories, and signals about internal states; the integration of contextual information with the actions and goals appropriate to that context; the choice among potential actions or goals, based on the predicted outcome of each possibility; and active maintenance of those choices or goals in memory, as a prospective code,* without completely dispensing with the alternatives*" (Genovesio & Wise, 2008, p. 101)*

1.2 Measuring Cognitive Flexibility

1.2.1 Cognitive Flexibility as a Multifaceted Construct

In accordance with its behavioral importance, cognitive flexibility has been measured in humans and other animals extensively (Bilalić, McLeod, & Gobet, 2008; Bonte, Kemp, & Fagot, 2014; Chrysikou & Weisberg, 2005; Davis, Vale, Schapiro, Lambeth, & Whiten, 2016; Diamond & Kirkham, 2005; Duncker & Lees, 1945; Hommel & Colzato, 2017; Ionescu, 2012; Kolodny, Edelman, & Lotem, 2015; Luchins, 1942; Meiran, 1996; Rumbaugh, 1971; Stoet & Snyder, 2003); yet, each of these studies employed entirely different paradigms. This is not to say that one task is better or more telling than the others; however, it is important to distinguish between the different types of cognitive flexibility that have (and are currently) being tested.

> *"And just as the sides of a cube are not cubes themselves but squares, it is hard to imagine that the sides of cognitive flexibility are cognitive flexibilities, too. Instead it might be more logical to consider these sides as essential mechanisms or processes (that can themselves be flexible) that contribute to the overall cognitive flexibility." (Ionescu, 2017, p. 6)*

1.2.2 Forced Switch Paradigms

The vast majority of cognitive flexibility metrics *require* subjects to switch strategies and then measure the ensuing deficits, typically in accuracy or response times (RTs). Among these forced switch tasks, the simplest paradigm is a reversal learning, where a previously valid strategy is no longer correct (Rosati, 2017). When the strategies are simple S-R associations (e.g., first the red circle is correct but later the blue circle is correct), this is referred to as a discrimination reversal task. Inflexibility or perseveration results when the subject continues to use their first-learned strategy after it is no longer correct and can be measured via a transfer index, calculated as performance on reversal trials divided by performance on the initial discrimination. Presumably the ability to extinguish the learned response determines how readily it is replaced (Roberts, 2008). In another forced switch paradigm, the A-not-B task, one of two targets is repeatedly baited and the subject is rewarded upon its selection. After an initial learning phase, the other location is baited. Subjects must switch from their learned response and instead select the newly baited location. Indeed, human infants (~9 months old) and monkeys continue to reach for the first-learned location (Diamond & Goldman-Rakic, 1989; Piaget, 1954; Zelazo, 2008).

Stroop tasks, in which subjects must ignore an attribute of a stimulus that conflicts with the correct response, are another measure of inhibitive ability and are also often reported alongside cognitive flexibility measures (Bunge & Wallis, 2008). Although inhibition is clearly involved in switching away from a learned solution, Friedman et al. (2008) recently proposed that inhibition may be an emergent property of other aspects of task shifting, such as updating (keeping a rule on-line in working memory) and switching (switching between rules), rather than a separable entity in and of itself (see Hommel & Colzato, 2017 for review).

Another commonly used forced switch measure of cognitive flexibility is the Wisconsin Card Sorting Task (WCST), or its simpler derivative the Dimensional Change Card Sort (DCCS) task (Ionescu, 2012; Jordan & Morton, 2012; Kirkham, Cruess, & Diamond, 2003; Manrique & Call, 2015; Zelazo et al., 2003). These measures require subjects to first sort a deck of multi-dimensional cards based on one stimulus attribute. For example, if the first rule is to sort based on color, cards with a red object would be placed in one pile and cards with a blue object would be placed in another, regardless of the objects' shapes. After this first sort, subjects are then asked to sort the cards based on the other dimension. For example, now subjects would need to place all the cards with a circle in one pile and all the cards with a square in the other pile, regardless of the objects' colors. Human children tend to perseverate with the first-learned rule; however, inflexibility on this task seemingly decreases with age (Kirkham et al., 2003; Zelazo, 2008). By 10 years old, children can perform similarly to adults (Chelune & Baer, 1986). Perseveration in adulthood is associated with executive impairment typical of schizophrenia and alcoholism (Sullivan et al., 1993) but RT deficits have also been observed in a typical adult population, suggesting that the 'attentional inertia' which leads to perseverative behaviors may negatively impact new strategy adoption throughout development (Diamond & Kirkham, 2005).

Deficits associated with switching response strategies, as opposed to repeating a strategy, are referred to as switch costs (Brass, Derrfuss, & von Cramon, 2008). Measured in response time delays or accuracy drops following a strategy change, switch costs are evident across a wide range of tasks and are often attributed to the cognitive reconfiguration allowing the new strategy to guide behavior (Lemaire, Luwel, & Brun, 2017; Luwel, Schillemans, Onghena, & Verschaffel, 2009; Meiran, 1996). Both children and adults exhibit switch costs (Ionescu, 2012; Zelazo, 2008) and they can be somewhat reduced if the subject knows that the switch is coming (Rogers & Monsell, 1995) or if they have more time to prepare for the new strategy (Arrington & Logan, 2004).

1.2.3 Optional Switch Paradigms

Problem solving strategies, or procedures for performing a task (Ardiale & Lemaire, 2012), generally fall into one of two categories. History-cued strategies are motivated by previous experience, while means-end strategies identify a path from the current state to a desired state; they are inspired by a goal (Sweller, Mawer, & Howe, 1982). Truly adaptive behavior goes beyond simply inhibiting and switching between known response contingencies. Often, one must engage in means-end problem analysis and actively select or even devise a solution strategy. This type of cognitive flexibility can be measured via insight problem solving tasks, wherein the solution requires approaching the problem in some novel way (Defeyter & German, 2003; Ionescu, 2012).

In the classic functional fixedness paradigm, an object for which the normal function is known must be used in an abnormal way to complete the task (Duncker & Lees, 1945; Knoblich, Ohlsson, & Raney, 2001). For example, when attempting to build a tower which needs to reach a certain height, subjects are less likely to succeed, which requires using a large box as the base of the tower, if the other building materials are inside of the box when provided to the subject. This mental ‘block,’ attributed to their prior knowledge regarding the box’s function, results in an inability to re-conceptualize the problem space (Duncker & Lees, 1945; Knoblich et al., 2001). In fact, difficulty replacing a familiar concept frequently affects humans’ performance on optional switch tasks, resulting in a ‘cognitive set.’

1.3 Cognitive Set: When Sufficiency Blocks Efficienc

1.3.1 What is Cognitive Set?

A cognitive set occurs when prior knowledge or experience biases the interpretation of task-relevant information, thus blocking a novel or better solution strategy. It has also been referred to as fixedness, einstellung, conservatism, proactive interference, negative transfer, entrenchment, satisficing, and task set (Badre, 2008; Bilalić et al., 2008; Bunge & Wallis, 2008; Dean, Kendal, Schapiro, Thierry, & Laland, 2012; Duncker & Lees, 1945; Hrubesch, Preuschoft, & van Schaik, 2009; Kolodny et al., 2015; Luchins, 1942; Sweller et al., 1982). Described as ‘pernicious’ (Bilalić et al., 2008), cognitive set has been documented in thousands of subjects across variable task designs (Bilalić et al., 2008; Chrysikou & Weisberg, 2005; Lemaire & Leclere, 2014; Luchins, 1942; Luchins & Luchins, 1950).

In 1942, Abraham Luchins published a detailed account of cognitive set, referring to it as the ‘Einstellung effect’ (Luchins, 1942). He used a ‘water jar’ task in which three jars were drawn and labeled with the quantity of water they contained. Using these values, subjects were required to add/subtract in order to ‘obtain’ a target quantity. The first five problems could all be solved via the same four-step strategy: Jar 2 - Jar 1 - Jar 3 - Jar 3 (see Figure 1.3-1).

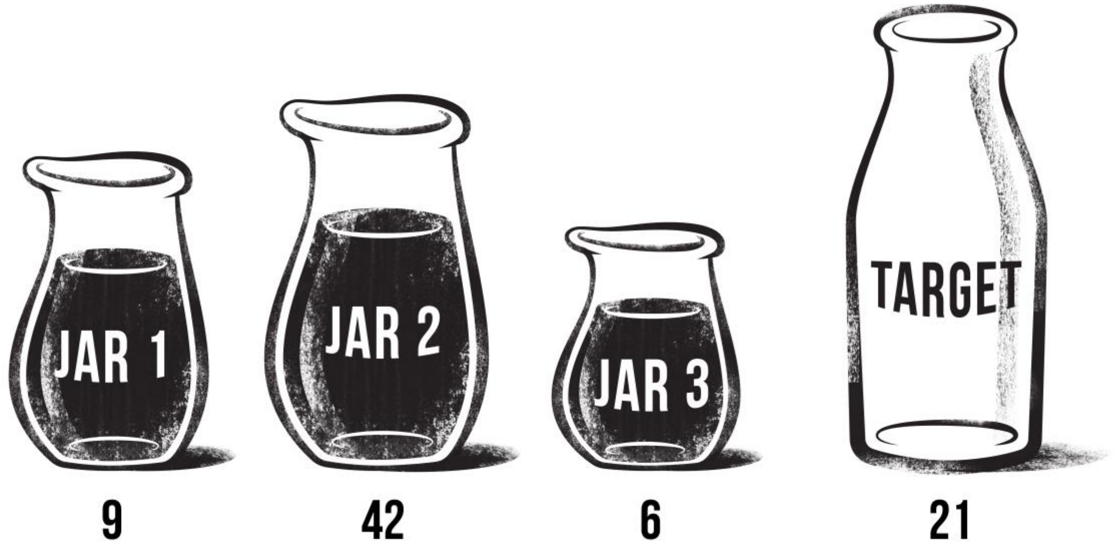


Figure 1.1 Example water jar problem, solvable via the four-step, learned rule.

However, the next two problems could be solved either by this four-step strategy *or* a simpler two-step strategy: Jar 1 - Jar 3 (see Figure 1.3-2).

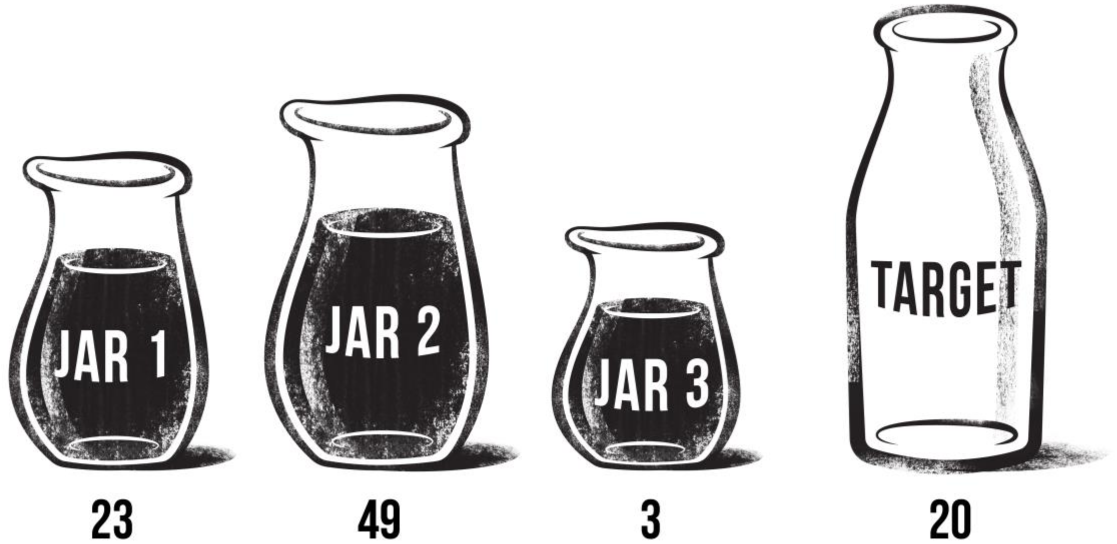


Figure 1.2 Example water jar problem, solvable via the learned rule or the better alternative

Luchins tested thousands of subjects of various ages and levels of education and found that the vast majority of subjects (70-100%, depending on the population sampled) continued to solve the problems using the more complicated, learned method rather than adopting a more efficient, but novel, two-step strategy.   In fact, even when subjects were given actual jars of water to manipulate (Luchins & Luchins, 1950), or the order of the jars was varied (Aftanas & Koppenaal, 1962), they persisted in their use of the learned strategy.

Luchins described this phenomenon as a *mechanization of problem-solving* and suggested that it arises in certain situations when previous experience with a problem-solving method infringes on a subject's ability to see the problem with an open state of mind (Luchins, 1942). For example, Bilalić et al. (2008) found that expert chess players, after finding one possible move, reported that they were looking for another. However, their gaze indicated that they continued to look at the solution they had already found. In another study, subjects were given a problematic example diagram of a bike rack and asked to redesign it such that these problems would be minimized; however subjects seemed unable to move past the examples and included even the problematic aspects in their designs (Chrysikou & Weisberg, 2005). Lemaire and Leclere (2014) noted that, when given two strategies for solving multiplication problems, subjects were more likely to repeat a strategy even if the other was more appropriate. Cognitive set emerges across a wide variety of tasks. But why? Why do subjects persistently use inefficient or problematic strategies when other, better strategies are available?

1.3.2 The Up-Side of Inefficiency

Cognitive set may not always be detrimental. Chunking, an enhanced association of certain behaviors involved in a response (Gobet et al., 2001), is thought to decrease cognitive load by automating certain behavioral responses (Kolodny et al., 2015; Smith & Graybiel, 2016). Further, while response-planning results in prolonged reaction times, especially at the beginning of a sequence (Rushworth, Croxson, Buckley, & Walton, 2008), routine behaviors can run on autopilot, they “require few of the processes needed for abstract response strategies” (Genovesio & Wise, 2008). Notably, switch costs are even apparent when strategy switching is voluntary and subjects tend to repeat a previous strategy more often than if their strategy was randomly selected (Arrington & Logan, 2004). In fact, increased speed, increased accuracy and decreased distractibility are some of the hallmarks of habit formation (Smith & Graybiel, 2016). Another consideration is that humans tend to avoid ambiguity, even sometimes choosing a riskier option rather than attempting one with an unknown outcome (Camerer & Weber, 1991).

Cognitive set affects a broad range of human problem solving conditions. Following a learned rule offers outcome security and can support the development of task proficiency; however, this eventually comes at the cost of task optimization. Mechanical solutions often occlude alternative, sometimes more efficient strategies, yet are seemingly pervasive in human problem solving.

1.4 The Evolution of Cognitive Flexibility

1.4.1 The Origins of Human Cognitive Flexibility

Currently, human culture is inseparable from technology. Over the past several millennia, humans have constructed an advanced assemblage of tools and techniques, a feat reliant upon the interplay between innovation and maintenance. Without invention, there is no novel technique to adopt and without behavioral maintenance, there is nothing to upgrade upon. Termed ‘ratcheting’ this ability to invent-adopt-upgrade has allowed humans to progress from basic stone tools to our current technological advancements (Davis et al., 2016; Hovers, 2012; Tennie, Call, & Tomasello, 2009; Tomasello, Kruger, & Ratner, 1993) and yet, ‘ratcheting’ is conspicuously less apparent in other primate species (Carr, Kendal, & Flynn, 2016; Davis et al., 2016). Thus, a perplexing contradiction arises: humans appear better equipped for cognitive flexibility than any other primate species; yet, when confronted with focused problem solving tasks, humans seem crippled by inflexibility. Elucidating when and how flexible strategy-use evolved and why similarly proficient invent-adopt-upgrade behavior appears to be lacking in nonhuman primates seems crucial for understanding these complexities.

Comparative cognitive research, especially across primate species, provides unique insights into the selection pressures potentially underlying the emergence of certain behaviors. Much of our evolutionary history has been shared with monkeys and, to an even greater extent, nonhuman apes (henceforth ‘apes’). Thus, comparing behaviors between monkeys, apes, and humans provides insights into which traits may have been present prior to their evolutionary divergences and which may have been uniquely derived (Rosati, 2017).

1.4.2 Cognitive Flexibility in Monkeys

Old world monkey (Cercopithecoidea) and ape (Hominoidea) lineages diverged approximately 25 million years ago (Stevens et al., 2013). Monkeys’ evolutionary proximity to humans and broadly similar neural organization has fueled numerous comparative studies aimed at identifying the origins and neural underpinnings of executive functions, including cognitive flexibility (Genovesio, Brasted, Mitz, & Wise, 2005; Genovesio & Wise, 2008; Manrique & Call, 2015; Rosati, 2017; Stoet & Snyder, 2008). Within these comparisons, several basic differences between monkeys and humans are apparent. First, monkeys often require extensive training to grasp response contingencies and even their most advanced socio-cognitive and executive processes are basic components of human cognition (Beran, Pate, Washburn, & Rumbaugh, 2004; Dean et al., 2012; Gallup, 1976; Manrique & Call, 2015; Stoet & Snyder, 2008; Visalberghi & Fragaszy, 2002). Second, monkeys' response times are recurrently faster than humans,' which may be attributable to faster neuronal conduction in smaller brains and/or some mechanical advantage (Ringo, Doty, Demeter, & Simard, 1994; Stoet & Snyder, 2008). Despite these differences, monkeys and humans show similarly inflexible behaviors on some tasks.

1.4.2.1 Evidence for Inflexibility in Monkeys

In basic forced-switch tasks, such as transfer index, which rewards subjects for selecting a certain response and then switches to rewarding a different response (Berg, 1948; Rumbaugh, 1971; Zelazo, Frye, & Rapus, 1996), monkeys and humans exhibit similar difficulties adopting the new solution (Bonté, 2011; Bonte et al., 2014; Ionescu, 2012; Jordan & Morton, 2012; Manrique & Call, 2015; Piaget, 1954; Rosati, 2017; Rumbaugh, 1971; Zelazo, 2008). In fact, on the Conceptual Set Shifting Task (CSST), a WCST derivative developed for nonhuman primates, both rhesus macaques and baboons increasingly produced perseverative behaviors with age, which is consistent with the human developmental trajectory on similar tasks (Bonté, 2011; Manrique & Call, 2015). Similarly, younger baboons outperformed older baboons on a transfer index task in both pre- and post-reversal sections (Bonte et al., 2014). Thus, monkeys and humans have shown similar inadequacies when switching behaviors; yet, when required to select between more abstract responses, several differences emerge.

1.4.2.2 Evidence for Flexibility in Monkeys

Like humans, monkeys are capable of learning abstract rules. In fact, rhesus macaques (*Macaca mulatta*) successfully enacted a conditional repeat-stay/change-shift strategy that requires completely ignoring S-R associations (Genovesio et al., 2005; Genovesio & Wise, 2008). In another study, a cue at the beginning of each trial signaled which of two strategies (e.g., judge color or judge orientation) rhesus monkeys and humans should apply to a stimulus (Stoet & Snyder, 2003). In some trials, both strategies called for the same response (e.g., choose *left*) but in others, the strategies were in conflict with one another (e.g., color strategy = choose *left*, orientation strategy = choose *right*). The monkeys successfully learned to adjust their strategies based on the cue and accurately responded to novel stimuli, confirming their grasp of the abstract contingencies (Stoet & Snyder, 2008). However, this study also revealed a key difference between monkeys' and humans' abstract rule-use. In stark contrast to humans, monkeys did not show slower response times when the strategy switched: in other words, they did not exhibit switch costs (Stoet & Snyder, 2003; however, see Huguet, Barbet, Belletier, Monteil, & Fagot, 2014 and Chapter 3 Footnote 1). As mentioned previously, humans are plagued by switch costs, which are even apparent after 23,000 trials (Stoet & Snyder, 2007), yet these monkeys were seemingly unaffected. Interestingly, the authors noted another difference between humans' and monkeys' task-switching abilities: unlike humans, monkeys showed lower accuracy and longer response times when facing the incongruent trials, those in which the two strategies signaled different responses (Stoet & Snyder, 2003). Thus, monkeys but not humans were significantly affected by irrelevant stimulus information.

> *"What humans are good at, compared with monkeys, is not switching between two tasks, but rather, locking on to a single task."* (Stoet & Snyder, 2003, p. 228)

Thus, although there is limited evidence for both flexible and inflexible forced switch strategy switching in monkeys, their propensity to forsake a learned strategy in order to adopt a better alternative, when it is available, is unclear.

1.4.3 Cognitive Flexibility in Apes

Within the Hominoidea superfamily, Hominidae or apes emerged ~13-18mya, with the *Pan-Homo* divergence occurring ~6-7mya (Perelman et al., 2011). As our closest extant relatives, apes provide an invaluable glimpse into the evolution of human cognitive processes. Apes possess relatively advanced cognitive abilities compared to monkeys and non-primate species (Biro & Matsuzawa, 1999; de Waal & Ferrari, 2010; Gallup, 1970; Haun & Call, 2008; Moore, 2016) but are outperformed by humans in many domains, especially with regards to language and social learning (Conway & Christiansen, 2001; Corballis, 2010; Heyes, 1993; Huber et al., 2009). It seems plausible that, given our shared ancestry, apes and humans might exhibit similar explorative/exploitative tendencies. However, if we consider the evolution of human culture to be predicated upon flexible strategy updating, it is also possible that a fundamental difference in explorative/exploitative behaviors gave rise to our vastly different technological trajectories.

To what extent apes are able to flexibly adopt new strategies is a topic of considerable debate and investigation (Hrubesch et al., 2009; Manrique, Völter, & Call, 2013; Marshall-Pescini & Whiten, 2008; Price, Lambeth, Schapiro, & Whiten, 2009; Van Leeuwen, Cronin, Schutte, Call, & Haun, 2013). In some measures, apes perseverate similarly to children (Beran, Washburn, & Rumbaugh, 2007; Rosati, 2017 for review; Rumbaugh, 1971; Vlamings, Hare, &

Call, 2010). Additionally, inflexibility was observed in all ape species on a reversal task, in which subjects first learned (over 100 trials) to push a lever one direction but then were required to extinguish this learned response and switch directions (Manrique & Call, 2015). Still other accounts state that, like humans, apes proficiently extinguish previous responses in A-not-B tasks (Amici, Aureli, & Call, 2008; MacLean et al., 2014), even performing equivalently to human children (Barth & Call, 2006). Yet the validity of these findings has recently been challenged. Recall that A-not-B tasks measure subjects' ability to switch from their learned (A) response and select the actual location of the food (B) but, typically, only three A responses are required before this 'strategy' is considered learned (Amici et al., 2008; Barth & Call, 2006; MacLean et al., 2014). Davis (2017) argued that three trials is not enough to make this claim and required subjects to utilize the A strategy a minimum of 20 times. Under these methods, chimpanzees responded at chance on the critical B trials, suggesting that their abilities to extinguish a learned response were, in fact, impaired.

Evidence regarding apes' abilities to switch between abstract rules is limited. One account suggested that on a DCCS task, chimpanzees' use of the second rule was impaired by their knowledge of the first (Moriguchi, Tanaka, & Itakura, 2011). However, the reported error rate post-switch was 52.9% and although the authors state that this was statistically different from chance, only 6 chimpanzees were tested and those only passed the pre-switch section in an average of 42.9% of sessions, shedding doubt on the extent to which even the first rule was adopted. If the first strategy was not established, difficulty learning the second strategy cannot confidently be attributed to cognitive set.

Importantly, chimpanzees are capable of abstract rule-use such as same/different judgments (Hopkins & Washburn, 2002; Nissen, Blum, & Blum, 1948; Oden, Thompson, &

Premack, 1988; Thompson, Oden, & Boysen, 1997), even using an exclusion tactic to group unknowns (Beran & Washburn, 2002). Further, in computerized mazes, chimpanzees are unimpaired when the correct path requires moving away from the goal briefly (Dolins, Schweller, & Milne, 2017; Fragaszy et al., 2009; Menzel & Menzel, 2007), suggesting that chimpanzees are capable of reasoning between and selecting appropriate strategies. However, to my knowledge, there have been no published investigations evincing chimpanzees' ability to switch between abstract solutions, prior to the current research.

Many social learning studies, investigating apes' ability to flexibly adjust their behavior after viewing a demonstration of an alternative method, have described both explorative and exploitative behaviors (Hopper, Kurtycz, Ross, & Bonnie, 2015; Yamamoto, Humle, & Tanaka, 2013). As this clearly relates to cognitive flexibility, I will briefly review their findings; however, with two caveats in mind. First, inflexibility in social learning contexts could also stem from an inability to extract the relevant information from a demonstration (for whatever reason), rather than from cognitive set (Dean et al., 2012). Second, inflexibility might also be observed if apes cannot perform or do not grasp the physical affordances of the alternative strategy. That being said, social learning paradigms are undoubtedly more representative of how group-specific behavioral repertoires (cultures) evolve or persist within ape populations.

1.4.3.1 Evidence for Behavioral Inflexibility in Apes

Several studies report that chimpanzees stick to a learned strategy, even after watching a demonstration of a more-rewarding alternative (Davis, 2017; Marshall-Pescini & Whiten, 2008). Marshall-Pescini and Whiten (2008) reported that among 5 chimpanzees, which were trained to use a dipping method to procure honey and then shown a more effective probing technique, all but one failed to switch to the new strategy. This led to speculation that chimpanzees may use a copy-when-dissatisfied heuristic, such that, as long as they receive *some* reward they do not switch to an alternative (Davis, 2017; Laland, 2004). Yet, without demonstrating that the probe technique would have been more readily adopted if they hadn’t first learned the dip method, inflexibility cannot be conclusively determined. A similar finding was observed in one population of wild chimpanzees, that were unable to use a stick to access honey, even when the stick was pre-placed inside of the honey-baited tube (Gruber, Muller, Reynolds, Wrangham, & Zuberbuhler, 2011).

Recently, Davis (2017) presented chimpanzees with a ‘pitfalls’ box, in which a reward could be accessed via multiple portals however, if chimpanzees used their learned method (Door 1), the best portion of the reward fell into a pit. Thus, to access the reward in its entirety, subjects needed to use a demonstrated, novel method (Door 2). The authors found that chimpanzees trained to use Door 1 adopted the Door 2 strategy after a median of 14 more attempts than naïve individuals, concluding that prior behavior credibly delayed but did not prevent use of the alternative.

In fact, some chimpanzees are conservative even when the alternative solution is easier. For example, chimpanzees proficient in a raking technique to obtain food from a food board, did not switch to a more efficient rattle method after their group members invented it (Hrubesch et al., 2009). In another example, after learning how to combine two components to construct a tool, chimpanzees continued to do so, even when the reward could easily be accessed via the individual components (i.e. did not require tool-construction; Price et al., 2009). This finding was mirrored in (Davis et al., 2016) who demonstrated that, while naïve chimpanzees easily adopted the most efficient solution strategy to a puzzlebox, 9 out of 11 chimpanzees initially trained to use a more complex strategy were not able to switch to the more efficient alternative following a demonstration. It is worth noting that for some of the tasks labeling apes as conservative, the alternative strategy is ‘locked’ during training (Davis, 2017; Manrique & Call, 2015). Subsequently, avoidance of that solution is not necessarily a function of the learned response – it could also be that subjects’ own personal unrewarding experience with that option delays (or prevents) their using it – compared to naïve controls that have no positive/negative experience with either solution.

Some of these findings are, indeed, reminiscent of the cognitive set literature in humans. If apes and humans similarly struggle to move past a known solution strategy to adopt a more efficient or more rewarding alternative, we might consider conservatism a homologous, ancestral trait. However, this assertion would be, as yet, unjustified.

1.4.3.2 Evidence for Behavioral Flexibility in Apes

In fact, another narrative heralds apes’ robust behavioral flexibility. This seems particularly evident when the alternative strategy yields a higher pay-off. For instance, in token exchange paradigms, chimpanzees flexibly switched to exchanging an unfamiliar token, when it became more highly rewarded than the familiar token (Van Leeuwen et al., 2013). Additionally, (Hopper et al., 2015) found that chimpanzees would flexibly travel to a more distant token-exchange location for a better payoff; echoing the potential foraging benefits of explorative behavior. Further, in a honey-baited tube experiment, a small subset of wild apes applied a familiar leaf-sponging technique to the novel task, suggesting a certain degree of innovation, marked by the ability to generalize a known strategy to a novel situation (Gruber, 2016).

Recently, (Davis, 2017) proposed that variable findings regarding cognitive flexibility in chimpanzees may be linked to the relative differences in strategy complexity between the learned solution and the alternative, such that relatively simple alternatives or modifications promote flexibility. This was supported by three lines of evidence (Davis, 2017). First, after learning how to slide a lever to achieve a peanut, chimpanzees were able to switch to another, more efficient strategy after it was demonstrated; however, they only switched if the other strategy yielded an increase in the reward value. Second, the authors introduced another reward retrieval paradigm and found that subjects would flexibly add a step to their response techniques, again, to attain a better reward. Third, when the difficulty of the first technique was increased, subjects readily switched to an easier alternative. These findings are in line with others demonstrating that simple strategy modifications are well within apes' capabilities (Lehner et al., 2011; Yamamoto et al., 2013).

Additionally, apes are certainly able to switch strategies when their learned method no longer works. For example, Manrique et al. (2013) showed that apes could readily master a puzzle box requiring consecutively more advanced solutions. Once a solution technique was learned, it subsequently became ineffective; thus, subjects were required to consecutively innovate other strategies. Further, disabled apes are certainly able to adapt familiar behaviors to deal with handicaps (Hobaiter & Byrne, 2010; Hockings et al., 2015), which means that, at least when it is required, apes are capable of flexibly modifying a strategy.

Further complicating the picture, chimpanzees have even exhibited flexible behavior above and beyond that of humans'. After being shown how to extract a reward from an opaque box, children and chimpanzees copied all demonstrated actions to achieve the reward. However, when the opaque box was replaced with a transparent one, children persisted in their use of all demonstrated actions, even the obviously unnecessary ones, such as tapping the outside of the box (Horner & Whiten, 2005; McGuigan, Whiten, Flynn, & Horner, 2007). Furthermore, this ‘overimitation’ was amplified in adults, who reproduced the unnecessary actions with even higher fidelity than children (McGuigan, Makinson, & Whiten, 2011). Yet, on the same task, chimpanzees discarded the causally irrelevant portions of the demonstration and accessed the reward with greater efficiency than either children or adults (Horner & Whiten, 2005). In fact, like humans, functional fixedness has even been suggested to afflict apes: presented with the floating peanut task, chimpanzees were unable to use water from their drinking source to float the peanut to an accessible height; however, when an alternate water source was provided, a small subset (5 out of 24) did succeed in procuring the peanut (Hanus, Mendes, Tennie, & Call, 2011).

1.4.4 Evolutionary Implications

Both flexible and inflexible behavior is observed in nonhuman primates. Seemingly, response style (exploitative vs explorative) can be influenced by a range of factors including the strength of the learned solution, the relative difficulty of enacting the alternative, and the reward difference between the two (Davis, 2017; Hopper et al., 2015; Marshall-Pescini & Whiten, 2008; Van Leeuwen et al., 2013).

Flexibility offers numerous advantages in terms of potential outcomes (Brosnan & Hopper, 2014; Sweller & Levine, 1982). More productive alternatives might be discovered or adopted. Novel resources might be exploited. In an ever-changing environment, it is nontrivial to navigate a novel situation successfully and this often requires innovative or explorative behavior (Holmes & Cohen, 2014). Adaptation is predicated upon flexible strategy use (Crone, Donohue, Honomichl, Wendelken, & Bunge, 2006; Lemaire & Leclere, 2014).

Yet, there are also potential disadvantages to flexibility. Explorative tendencies may predispose an individual towards distractibility (Hommel & Colzato, 2017; Stoet & Snyder, 2008). Effort spent developing a new technique or even the cognitive processes associated with switching between techniques, might be costly (Holmes & Cohen, 2014; Stoet & Snyder, 2003). Moreover, environmental change could be fleeting, rendering an adaptation useless (Brosnan & Hopper, 2014) or at the very least, a waste of time. Further, an innovated strategy is untested, whilst the pre-existing behavior sufficiently supported survival thus far (van Schaik et al., 2016). In other words, if the current solution elicits some reward, the relative benefit of exploring alternatives might be negligible or even detrimental. Hence the colloquialism: “If it ain’t broken, don’t fix it.”

Yet, conservatism, especially in regards to abstract rule-use, is restricting. It impedes innovatory behavior thereby lowering the probability of discovering adaptive solutions or niches (Brosnan & Hopper, 2014). Moreover, without the flexibility to choose between alternatives, optimization is unlikely.

1.5 Cross-Cultural Influences on Cognitive Flexibility

Thus far, I have compared humans’ cognitive flexibility with monkeys’ and apes,’ without fully acknowledging that there is much evidence showing inter-individual variability in humans’ explorative/exploitative tendencies. Here, I will introduce potentially mediating factors, specifically focusing on cross-cultural influences, which might affect humans' propensities to break cognitive set.

1.5.1 Humans and Rule-Use

Human problem solving is unequivocally advanced. To a large extent, this dexterity is facilitated by codified rules-of-thumb and equations, which enable us to accurately solve many problems, even without necessarily understanding *why* a solution strategy is effective. For example, one can successfully calculate the circumference of a circle without ever grasping the intricacies of pi. Undoubtedly, codified solution strategies offer a tremendous advantage when solving many, similar problems. For this reason, repetitive rule-use is a staple of Western education, yet the real-world applications of this approach are dubious. First, how often do every day problems require identical solution strategies? Simply determining when and where to cross the road affords infinite solutions and reproducing the actions that allowed you to safely cross previously could be morbidly ineffective under the current conditions. Second, might rote learning reasonably lead to cognitive inflexibility (Star & Seifert, 2006)? In a novel problem space, the search for a solution takes precedence but if a solution is provided and subsequently memorized, search might never truly occur. With regards to cognitive set, perhaps alternative solutions are not found because they are not sought. In fact, stoked by partial differences in alternative strategy-use between publicly and privately schooled children, Luchins (1942) repeatedly professed that typical Western educational practices might be largely accountable for the observed conservatism.

1.5.2 Purported Universality of Cognition

Thousands of subject have been tested on Luchins' water jar task (see Section 1.3.1) under various manipulations and a consistent majority persist in their use of the learned rule. This has led to a conclusion that, within human problem solving, cognitive set is universal (Aftanas & Koppenaal, 1962; Luchins, 1942; Luchins & Luchins, 1950; McKelvie, 1984). Yet, despite attempts to account for subjects' age, occupation, and education, only Western (American, British, and Canadian) subjects were tested, leaving the potential for cross-cultural differences in susceptibility to cognitive set completely unexplored. Recently, investigators have rightfully called into question the 'universality' of many cognitive processes, of which our understanding is predominantly derived from Western, Educated, Industrialized, Rich, Democratic (WEIRD) subjects' data (Clegg & Legare, 2016; Henrich, Heine, & Norenzayan, 2010; Legare & Nielsen, 2015). One particularly concerning assessment found that in 2007, 67% of studies in the *Journal of Personality and Social Psychology*, solely collected data from undergraduate psychology students (Arnett, 2008; Jahoda, 2016). In fact, studies that defy this trend often find that WEIRD data are at the extremes of more globally diverse samples. Which is not surprising, given Westerners' highly unusual social and ecological environments compared to those of the vast majority of humans currently, and those in which all humans belonged to very recently in our evolutionary past (see Henrich et al., 2010 for review). The comparatively few studies investigating cross-cultural cognition have identified differences in both perceptual and conceptual domains.

1.5.3 Westerners' Global Perceptual Bias

Visual assessment of one's surroundings is crucial to survival; yet the specific demands are necessarily different in different environments. What may be visually important to a hunter (e.g., judging the speed and directionality of movement) will not be the same as what is important to an artist (e.g., attending to the minute details within a visual scene). Thus, unsurprisingly, humans have been shown to exhibit differences in visual perceptual processing.

Westerners have been found to 1) make more accurate judgments regarding a global figure comprised of conflicting local figures than vice versa and b) preferentially group figures based on global rather than local features (Caparos, Ahmed, et al., 2012; Davidoff, Fonteneau, & Fagot, 2008; Navon, 1977). In the visual system, the lateral geniculate nucleus, which receives visual input directly from the retina and projects to the primary visual cortex, is comprised of six laminae: four layers of cells with small receptive fields (parvocellular) and two layers of cells with large receptive fields (magnocelulllar). Indeed, the interplay between magno- and parvocellular inputs appears to be the source of humans' observed global bias, evidenced by the ability to attenuate it by suppressing a portion of the cells in the magnocellular pathway (Michimata, Okubo, & Mugishima, 1999). Thus, one might expect global/local perceptual biases to vary consistently between individuals.

1.5.4 Cross-Cultural Differences in Perceptual Biases

Differences in local/global perceptual biases have been reported across cultural, ethnic, and even religious groups (see Hommel & Colzato, 2017 for review). For example, in a task which requires subjects to judge the verticality of a line (rod) within a frame, East Asian participants were more affected by the orientation of the frame than Americans (Ji, Peng, &

Nisbett, 2000). In a similar task, Japanese participants more accurately reproduced a line that was proportional to the frame, compared to Americans who were better able to ignore the frame and reproduce the absolute size of the line; this was also true for Americans living in Japan and Japanese living in America (Kitayama, Duffy, Kawamura, & Larsen, 2003). Eye tracking analyses showed that, when viewing a picture, American participants fixated more on the focal object, but Chinese participants tended to look at the background (Chua, Boland, & Nisbett, 2005). Additionally, this attentional selectivity seemed to impact information processing: when describing a scene, Japanese participants mentioned the background first, but Americans mentioned foreground objects first (Masuda & Nisbett, 2001), which influenced their ability to recognize the scenes after changes to either the foreground or background were made (Masuda & Nisbett, 2006). However, similar to Americans, on the hierarchical figures task (Navon, 1977), Japanese and British participants' exhibited equivalently global biases (Caparos, Ahmed, et al., 2012).

Humans have also been shown to possess local perceptual biases. In direct comparisons, the seminomadic Himba of northern Namibia exhibited a 'dramatic' local bias on hierarchical figure tasks compared to Western participants (Davidoff et al., 2008), and remarkably, this bias decreased as the number of exposures to an urban environment increased (Caparos, Ahmed, et al., 2012). Further, compared to British participants, Himba were significantly less affected by the Ebbinghaus illusion, in which the size of an inner object appears to change depending on the composition of surrounding objects (Caparos, Ahmed, et al., 2012; de Fockert, Davidoff, Fagot, Parron, & Goldstein, 2007). Likewise, in a categorization task, the Himba grouped shapes with local similarities even when they fell into different Western shape categories (i.e. triangles, circles, and squares; Roberson, Davidoff, & Shapiro, 2002). Lastly, the Himba seem to possess greater attentional control than Westerners, as evidenced by their enhanced performance on an Eriksen-type flanker task compared to British participants (de Fockert, Caparos, Linnell, & Davidoff, 2011). Thus, there is clear evidence that visual perception differs between cultures. However, if and to what extent this interacts with environmental factors to affect problem solving is less clear.

1.5.5 Cross Cultural Differences in Problem Conceptualization

The ways in which humans conceptualize problems and their solutions, differ substantially across human populations. Many environmental factors may contribute to humans' performance on cognitive tasks. For one, cross-cultural differences in linguistic encoding might lead to variation in subjects' abilities to remember and discriminate between stimuli, especially those which fall along a continuum such as color (Davidoff, Davies, & Roberson, 1999; Davidoff & Fagot, 2010; Roberson, Davidoff, Davies, & Shapiro, 2005) or quantity (Frank, Everett, Fedorenko, & Gibson, 2008). That said, Ji, Zhang, and Nisbett (2004) found that culture but not language influenced subjects' categorization: when provided with three words and instructed to select the two most related, Americans preferred to group the words based on categorizations (e.g., *monkey* and *panda*) whereas Chinese participants preferred relationship-based categorizations (e.g., *monkey* and *banana*), regardless of testing language. Similarly, Chinese children grouped items based on functionality while American children utilized more categorical groupings (Chiu, 1972; Nisbett, Choi, Peng, & Norenzayan, 2001) and Asian adults were more likely to interpret events as stemming from environmental causes, compared to Americans who more often attributed them to internal causes (Norenzayan & Nisbett, 2000). There is even some evidence that problem-solving approach might differ across cultures.

Chinese and Korean subjects evoked more intuitive (or holistic) reasoning styles, while Americans (even Asian Americans) relied on more analytic problem solving approaches (Norenzayan, Smith, Kim, & Nisbett, 2002).

However, some studies have noted similarly inflexible problem solving approaches between cultures. For example, overimitation, wherein subjects copy even irrelevant portions of a demonstration, has been observed across Western and remote cultures alike (Horner & Whiten, 2005; McGuigan et al., 2007; Nielsen, Mushin, Tomaselli, & Whiten, 2014; Nielsen & Tomaselli, 2010; however, see Berl, 2015). Yet, another set of studies revealed that overimitation varied within individuals across cultures, depending on context: Clegg and Legare (2016) found that both American and Vanuatuan children more closely imitated instrumental (i.e., object-related) compared to conventional (i.e., socially determined) demonstrations (see Legare & Nielsen, 2015 for review of cultural influences on learning). Additionally, similar to Westerners, adolescent Shuar of Equadorian Amazonia were shown to be affected by functional fixedness during problem solving (German & Barrett, 2005). However, in typical measures of cognitive flexibility such as the WCST, some cross-cultural differences have been found (Avila, 2013; Shan, Chen, Lee, & Su, 2008), yet in one case, these were attributed to flaws in the task itself rather than potential variation in cognitive processes across cultures (Coffey, Marmol, Schock, & Adams, 2005).

1.5.6 Impact of Cultural Differences on Problem Solving

Several hypotheses have emerged, which attempt to assimilate the observed cross-cultural differences in perceptual bias (global/local), attentional bias (context/focal), and response style (flexible/persistent) in both perceptual processing and problem conceptualization across cultures.

For instance, the social structure hypothesis heralds the impact of subjects' cultural interdependency on cognition. Under this hypothesis, societies that revere the collective (the common example is East Asian) rather than the individual (the common example is American) are predicted to exhibit global rather than local perceptual biases, which in turn are predicted to support holistic rather than analytic response styles (Nisbett et al., 2001). Although one study found that priming subjects with more interdependent pronouns (e.g., we, our, us) elicited enhanced performance on global responses and vice versa for independent pronouns (e.g., I, me, mine) and local responses. (Kuhnen & Oyserman, 2002), there is substantial evidence to discredit this hypothesis.

First, not all interdependent peoples exhibit global perceptual biases. As mentioned in Section 1.5.4, Himba people, who live in extremely interdependent groups, exhibit profoundly *local* perceptual biases (Davidoff et al., 2008; Roberson et al., 2002). Second, some of these findings are not reproducible. A recent study (N = 363) showed that Japanese participants were instead *less* globally oriented than American or Argentinian participants and this was reconfirmed (N = 1,843) with a representative sample of Japanese and American participants aged 20-69 (Oishi et al., 2014).

Under the visual clutter hypothesis, the characteristics of a culture's typical visual environment are said to contribute to local/global perceptual biases. In other words, global perceptual bias is thought to develop from regularly encountering cluttered visual scenes; whereas local perceptual biases might result from sparse visual scenes (Caparos, Ahmed, et al., 2012; de Fockert et al., 2011). Both Japanese and American participants, primed with Japanese scenes (which were found to be more ambiguous and contain more elements than American scenes) attended more to context than those primed with American scenes (Miyamoto, Nisbett, & Masuda, 2006); however, this seems to suggest instead that attentional biases shift based on recent contextual information. In fact, Linnell, Caparos, de Fockert, and Davidoff (2013) demonstrated that attentional focus was suppressed in urban compared to traditionally-living Himba and de Fockert et al. (2011) showed that, in stark contrast to Westerners, Himba participants were not even distracted by a moving singleton. Thus, there is a potential link between visual clutter and attentional selectivity. However to what extent attentional selectivity and global/local perceptual biases are conflated is unclear (Caparos, Linnell, Bremner, de Fockert, & Davidoff, 2012). Moreover, how visual clutter may influence problem solving approach (Hommel & Colzato, 2017), is entirely speculative. When placing Kitayama et al.'s (2003) rod and frame study into context with other cross-cultural perceptual findings, Hommel and Colzato (2017, p. 5) wrote:

> *"With respect to control styles, this can be taken to imply that US-Americans have a stronger bias towards persistence (the control style inducing an analytic processing mode) than Japanese have and/or Japanese participants have a stronger bias towards flexibility (the control style inducing a holistic processing mode) than US-Americans have."*

Given the diversity of visual environments and doctrine to which individuals within many of the tested communities are exposed, I would interpret such group differences as evidence that these are *not* influential predictors of cognitive flexibility. Furthermore, when the same individuals ($N = 70$) completed two different metrics of wholistic/analytic response style, the two scores did not significantly correlate, indicating that response style cannot be simplified into binary wholistic/analytic mindsets (Peterson & Deary, 2006).

In summary, even a rudimentary understanding of the impacts of culture on cognitive flexibility is currently lacking. The existing hypotheses are confounded with differences in perceptual biases and interpersonal belief systems, which appear related to problem solving approach, at best, tangentially (Hommel & Colzato, 2017; Nisbett et al., 2001; Nisbett & Miyamoto, 2005).

1.6 Current Research Aims and General Approach

1.6.1 Specific Aims

In an effort to explore *how* and *why* susceptibility to cognitive set might differ between primate species and across human cultures the following experiments aimed to:

1) Identify baboons', chimpanzees', and humans' relative propensities toward cognitive set during abstract problem solving and, by comparing alternative strategy use across these species, provide a putative evolutionary trajectory for how it may have evolved within the primate lineage.
2) Identify the relative impacts of perceptual and conceptual influences on humans' propensity toward cognitive set by determining:
    i. Whether or not humans a) do not *see* alternative strategies or b) do not *understand* that they can use them.
    ii. The relative impacts of a) global/local perceptual biases and b) cross-cultural differences on alternative strategy-use.

1.6.2 General Approach

For these experiments, I devised and used a nonverbal, non-arithmetic, computerized measure of susceptibility to cognitive set: the Learned Strategy – Direct Strategy (LS-DS) task.

For each LS-DS trial, a demonstration reveals which of the four locations constitutes Square1 (flashes red first) and which constitutes Square2 (flashes red second). Throughout several levels of training, subjects learn to utilize a three-step sequence, in which they first copy the demonstration by selecting Square1 and Square2, and then select a Triangle which appears in one of the two remaining locations. This three-step sequence, (Square1→Square2→Triangle) constitutes the learned strategy (**LS**; see Figure 1.6-1a). After training, once subjects consistently utilize the LS, experimental trials are presented, in which baseline (BASE) and probe (PROBE) trials are interspersed. For BASE trials, the Triangle is *hidden* (not visible) until Square1 and Square2 have been accurately selected. For PROBE trials, importantly, the Triangle is present from the beginning of the demonstration and remains on the screen throughout subjects' response. Thus, subjects can use the LS (See Figure 1.6-1b) *or* they can use a direct strategy (**DS** or the shortcut) by skipping Square1 + Square2 and simply selecting the Triangle immediately (See Figure 1.6-1d). Additionally, the task permits a third strategy, termed the switch strategy (**SS**), wherein subjects initially select Square1 but then skip Square2 and select the Triangle; however this was not a common approach (Figure 1.6-1c).

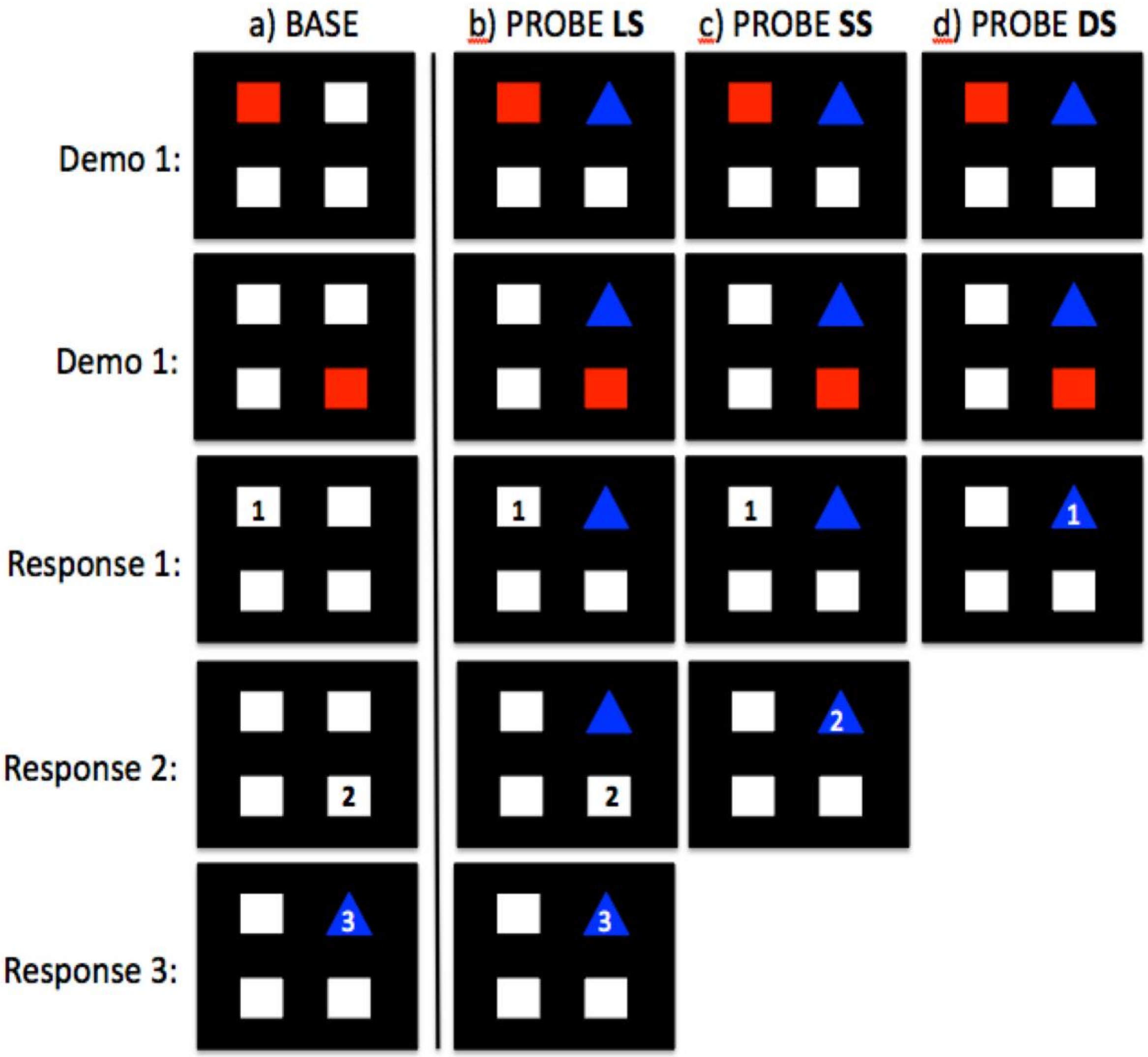


Figure 1.3 An exemplar LS-DS trial depicting the Top left → Bottom right → Top right configuration. a) On BASE trials, only the LS can be successfully employed because the Triangle is 'hidden' until Square1 and Square2 are correctly selected. On PROBE trials, subjects may utilize any of three response strategies to be correct: b) the learned strategy (LS), c) the switch strategy, (SS), or d) the direct strategy (DS). Numbers indicate the correct Response 1, Response 2, and Response 3, for this configuration.

1.6.3 Overview of Research Chapters

In Chapter 2, baboons' and humans' ability to forego the LS and adopt the DS is compared. This chapter is presented as it was originally published (Pope, Meguerditchian, Hopkins, & Fagot, 2015)*;* however small changes were made regarding the way in which DS-use was calculated between this and all subsequent chapters (See Chapter 3 Footnote 2). In Chapter 3, chimpanzees' DS-use is investigated and compared with baboons' and humans' data from Chapter 2; the evolutionary implications for cognitive set are then discussed. Chapter 4 presents data regarding humans' ability to see the alternative and how demonstrating the shortcut affects its use. In Chapter 5, Americans' DS-use is compared to that of the semi-nomadic Himba of northern Namibia, to investigate the relative influences of global/local perceptual biases and culture on cognitive set.

1.6.4 Overarching Hypotheses

I hypothesize that:

I) Cognitive set differentially affects humans, chimpanzees, and baboons.

II) Conceptual understanding of the alternative strategy as a viable option, rather than the ability to perceive the alternative, determines subjects' susceptibility to cognitive set.

Evidence supporting Hypothesis I is presented in Chapters 2 and 3, which report that, compared to baboons and chimpanzees, humans were significantly better able to learn the LS, were more affected by switch costs when shifting between the LS and DS, and were more susceptible to cognitive set on the LS-DS task. I posit that the ability to verbally encode abstract rules results in their being more firmly represented (i.e., less easily replaced), which contributes to humans' elevated susceptibility to cognitive set. However, alternative interpretations are discussed in Chapters 2, 3 & 6.

Chapters 4 and 5 provide support for Hypothesis II. In Chapter 4, subjects' gaze indicated that they did, in fact, see the alternative strategy but did not use it. Further, their ability to use the shortcut was, not surprisingly, enhanced after viewing a video demonstrating the DS (compared to controls who watched a video demonstrating the LS). In Chapter 5, Himba subjects were shown to be less susceptible to cognitive set, compared to Americans; however, this did not co-vary with known group differences in global/local perceptual processing. Further, only American subjects benefitted from a prompt aimed at emboldening subjects to try new things. I posit that cultural differences in educational background might contribute to the observed differences in flexible strategy-use. Specifically, I suggest that the mechanized rule-use typical of Western education might promote cognitive set, by predisposing subjects to respond by using learned strategy.

1.6.5 Summary and Significance

Appropriately replacing sufficiency with efficiency seems crucial to adaptive behavior. Likely, as with many cognitive processes, individuals are capable of operating within a range of cognitive flexibility, such that the persistence of any given behavior is context dependent. The LS-DS measures a single context: when multiple abstract solutions (LS, SS, and DS) are present, which vary only in efficiency. Thus, although I will sometimes refer to cognitive flexibility and shortcut-use synonymously, to be clear, I consider cognitive set as merely one facet of cognitive flexibility.

The results of this research discern between whether set arises from a difficulty in recognizing alternative strategies or a difficulty in executing them. Additionally, it identifies when human-like propensity toward cognitive set might have arisen within the primate lineage. It is clear that mechanized rule-use can be beneficial and is often adaptive in problem solving (Langer & Lois, 1979; Sweller & Gee, 1978). However, understanding the mechanisms and evolution of alternative strategy use on the LS-DS task may provide a foundation for future studies aimed at enabling participants to consciously break set, when it is beneficial.

2 BABOONS (*PAPIO PAPIO*), BUT NOT HUMANS, BREAK COGNITIVE SET IN A VISUOMOTOR TASK

A version of this manuscript was published in *Animal Cognition*, the final version is available at https://link.springer.com/article/10.1007/s10071-015-0904-y

Sarah M. Pope[1,2], Adrien Meguerditchian[2,3], William D. Hopkins[1,4], and Joël Fagot[2,3*]

[1] Neuroscience Institute and Language Research Center, Georgia State University, Atlanta, Georgia, USA

[2] Aix-Marseille University, Laboratory of Cognitive Psychology, UMR7290 CNRS and Federation 3C, Marseille, France

[3] Brain and Language Research Institute, Marseille, France Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, Atlanta, Georgia, USA

Abstract

Through codified rule-use humans are able to accurately solve many problems; however, mechanized strategy-use can also be costly. Cognitive set occurs when a familiar solution strategy blocks a better alternative. Despite half a century of research on cognitive set in humans, there have been no attempts to investigate whether it impacts nonhuman species' abilities to solve abstract problems. The current study utilized a non-verbal, computer task to compare cognitive set between 104 humans and 15 baboons (*Papio papio*). A remarkable difference was found between humans' and baboons' abilities to break cognitive set. Consistent with previous studies, the majority of humans were highly impaired by set, yet baboons were almost completely unaffected. Analysis of the human data revealed that children (ages 7-10) were significantly better able to break set than adolescents (11-18) and adults (19-68). Both the evolutionary and developmental implications of these findings are discussed.

2.1 Introduction

As problems increase in complexity, so too do their solutions. To mediate the difficulties of solving these complex problems, rules may be established which give the correct answer yet bypass problem reanalysis. Rules allow many similar problems to be solved efficiently and are often used to teach problem-solving skills (Henderson & Pingry, 1953; Rohrer & Taylor, 2006). However, problem solving by rule-use is not solely beneficial, as is the case when a learned strategy is less efficient than an alternative. Cognitive set, also termed ‘einstellung’ or ‘mental set,’ occurs when a subject successfully learns a rule to solve several problems but is unable to switch to a more efficient method when it becomes available (Luchins, 1942; Ruscio & Amabile, 1999). Thus, an inability to break cognitive set occurs when that learned solution occludes other problem-solving methods. In other words, once a rule is adopted, other options are not explored. An inability to break cognitive set leads to inefficient problem solving by preventing the use of alternative, sometimes better, problem-solving methods.

In 1942, Abraham Luchins showed that thousands of humans struggled to break cognitive set in order to use a more efficient solution. Luchins’ (1942) task required participants to obtain a target quantity by adding and subtracting three given values. The first five of these problems could be solved by a single, somewhat complex, rule. However, these were followed by two hybrid problems, solvable both by the learned rule and a more efficient, direct method. Luchins (1942) found that 70-100% of subjects persisted in using the learned rule rather than switching to the direct method. However, despite its pervasiveness, the underlying causes of humans’ susceptibility to cognitive set and potential methods to combat it are opaque.

Several factors influence, to varying degrees, subjects’ abilities to break a mental set including instruction (Aftanas & Koppenaal, 1962; Luchins & Luchins, 1950), working memory availability (Beilock & Decaro, 2007), speed requirements (Luchins, 1942), amount of training (Crooks & McNeil, 2009; Luchins, 1942), and similarity between problems (Sweller et al., 1982). Further, although Luchins (1942) reported no substantial age effects on cognitive set, age has been shown to affect other similar types of problem solving. 'Functional fixedness' is described as occurring when an object's use as a tool is dramatically hindered by a subject's experience with it in another functional role (Duncker & Lees, 1945). In a tool-use task, Defeyter and German (2003) reported that five-year-old children were unaffected by their previous experience with a tool, yet seven-year-olds and adults easily fell victim to functional fixedness. Despite these accounts, little is known of the differences in cognitive set between children, adolescents, and adults.

Understanding the evolutionary origins of cognitive set may aid in understanding it's pervasiveness in human problem solving. However, previous research on how the mechanization of set might have evolved is nonexistent. This is likely due to the impossibility of a comparative analysis using Luchins' task, which used an arithmetic problem. Studies comparing adults and nonhuman primates using computer paradigms have noted differences in problem-solving performance that may be relevant to set-breaking behavior. First, differences in sequential processing have been reported between nonhuman primates and adults. Ohshiba (1997) noted that macaques' response times to a simultaneous chaining task increased as they progress through the sequence, indicating that they are using a 'serial search strategy.' Conversely, human adults' responded slowly to the first item in the sequence but quickly to the rest of the items, indicating that they were using a 'collective search strategy' and were mentally identifying the entire sequence before reproducing it (Conway & Christiansen, 2001; however, see Fagot & De Lillo, 2011 and Beran et al. 2004). These results may be applicable to problem-solving in general, with macaques operating in a more local manner (each step is independent) and humans in a more global one (each step is part of the sequence). This is in line with findings comparing perceptual biases between humans and baboons *(Papio papio*), another old world monkey species. Baboons were found to respond more quickly to local stimuli while human adults responded more quickly to global stimuli (Deruelle & Fagot, 1998). If we consider that the key to avoiding cognitive set is likely rooted in an ability to see and utilize the individual steps within a rule, humans' collective approach may be what is driving their inability to break set. Further, Stoet and Snyder found that macaques' problem solving was more affected by distractions than human adults' (2003), suggesting that they may be less focused on the problem-solving rule and more attentive to individual variation between problems. Indeed, this may provide old world monkeys with an increased awareness of the alternative method in a cognitive set task.

The current research has two main goals. First, it investigated the evolutionary origins of cognitive set by comparing humans to baboons in a computerized, nonmathematical cognitive set task. We hypothesized that the ability to break set would be different between the two species due to the differences in perceptual and sequential processing between old world monkeys and humans. Indeed, extreme differences were found between the two species. Baboons were almost entirely immune to the effects of set, while the majority of humans did not break away from the learned rule. Second, the developmental trajectory of cognitive set in humans was analyzed by comparing children, adolescents, and adults. Children were significantly more likely to break cognitive set than either adolescents or adults. These findings are discussed from both evolutionary and developmental perspectives.

2.2 Methods

2.2.1 Subjects and Materials

Baboon data were collected from 15 subjects (ages 1.8-9.3 years), including six males (*Mean* age = 5.3, *SD* = 2.68) and nine females (*Mean* age = 5.1, *SD* = 2.36), living in a larger social group of 24 individuals located at the CNRS "Station de Primatologie", Rousset-sur-Arc, France. Baboons were tested via 10 automated learning devices for monkeys (ALDMs; Fagot & Bonte, 2010; Fagot & Paleressompoulle, 2009), which were directly attached to an outside 700 m^2 enclosure. Subjects had unrestricted access to the ALDMs which consisted of a 70 cm × 70 cm × 80 cm testing chamber with a view port and two hand ports. The view port allowed subjects to see the 19-inch LCD touchscreen monitor (1939L Open-Frame Touchmonitor, Elo Touch Solutions). As subjects reached through the hand ports, a microchip was read for subject identification, which prompted the program to resume the trial list at the appropriate place for that subject. For correct responses, the ALDMs automatically dispensed several grains of dry wheat. The experiment was programmed using EPrime (Version 2.0, Psychology Software Tools, Pittsburgh). The local "Provence Alpes Côte d'Azur" ethic committee for experimental animal research approved the use of the ALDM procedure.

Human data were collected from 104 subjects (ages 7-68), including 40 males (*Mean* age = 26.85, *SD* = 17.7) and 64 females (*Mean* age = 25, *SD* = 17.7). Subjects were recruited via a sign which read "Would you like to be a part of a scientific study?" and tested at Zoo Atlanta, in Georgia, USA. Humans were tested behind a curtain in a 'booth' along a main path at Zoo Atlanta on a 19-inch LCD touchscreen monitor (1915L Desktop Touchmonitor, Elo Touch Solutions). The experimenter was nearby, but separated from the subject by a curtain and remained inattentive. Additionally, family members often remained in the general vicinity but were asked to remain inattentive and out of sight of the participant. Participants were given headphones (Koss On-Ear KPH Headphones, KPH7W) to hear sounds elicited by incorrect or correct responses. Correct responses were followed by a cartoon of a present, which increased in size with each correct response. After the fourth correct response, subjects were allowed to choose a sticker and the present size was reset. The same instructions were given to both children and adults explaining the correct/incorrect response screens and stated that they "would need to touch the shapes to figure out the right answer." Subjects could choose to stop testing at any point and 27 (6 adults, 6 adolescents, and 15 children; 20.6% of total attempts) either stopped of their own volition or were casually stopped (by displaying the end screen) if they had not passed the training within 15 minutes. All human methods were approved by the Zoo Atlanta Research Committee and the Georgia State University Institutional Review Board prior to testing.

2.2.2 Testing

The testing phase consisted of two conditions, baseline (BASE) and probe (PROBE). Trials began after the subject touched a fixation cross. First, two demonstration slides (150 ms each) displayed a 300 x 300 pixels red square in one location and then again in another location (out of four possible locations) on a light blue background (Fig 2.2-1). Subjects were then given a response screen, consisting of four white squares in the four available locations on a dark blue background. For correct responses, subjects were required to reproduce the demonstration by touching the two white squares located where the red squares had been in the demonstration, in the correct order. If subjects touched an incorrect square they were shown a 3 second 'time-out' screen before the next trial was cued. In the BASE condition, if subjects accurately reproduced the demonstration, they were presented with a blue triangle in place of one of the two remaining white squares (Fig 2.2-1a). To answer correctly, subjects needed to touch the blue Triangle. This sequence (touch Square1, touch Square2, touch Triangle) constituted the learned strategy (LS). However, in the PROBE condition, the blue Triangle was shown throughout the two demonstration squares and remained visible in the same location on the response screen (Fig 2.2-1b). Therefore, subjects could continue to use the LS, but were also able to touch the Triangle directly and receive a more immediate reward. This more efficient response (touch Triangle) constituted the direct strategy (DS). Importantly, in BASE trials, the Triangle was revealed after the first two correct responses. However, if subjects happened to touch where the Triangle was located during any of the response screens, they were rewarded, thereby enabling us to measure the number of times subjects might accidentally touch the Triangle's location regardless of it being visible.

Baboons were given 720 testing trials (576 BASE and 144 PROBE), humans 11 years and older were given 96 testing trials (48 BASE and 48 PROBE) and humans under 11 were given 48 testing trials (24 BASE and 24 PROBE). The number of trials presented to humans was reduced to minimize fatigue (as they had to complete the experiment in one sitting). It should be recognized that for baboons, the ratio of BASE to PROBE was 4:1, while for humans it was 1:1. The 4:1 ratio promotes the use of the LS in baboons. This is due to BASE trials not being readily solvable with the DS method, as the Triangle is hidden. Therefore, more BASE trials means more LS use. After collecting pilot data indicating that humans preferred the LS even with the 1:1 ratio, we opted to collect an even number of PROBE and BASE trials for humans to minimize the duration of the test sessions, while keeping the number of test trials large enough for statistical analyses and cross-species comparison.

2.2.3 Training

The three training procedures were carried out over the course of 14 days for baboons and immediately preceding the testing phase for humans. Baboons completed an average of 12,945 training trials (*SD* = 4346), while humans completed an average of 35.2 trials (*SD* = 18). One of the baboons did not pass Training level 2 but successfully passed the more difficult Training level 3; thus, his data were included.

For baboons, Training 1 consisted of 96-trial blocks. Trials were randomly selected from 24 possible square/triangle configurations. Each trial began with a fixation cross, followed by two demonstration slides (each 150 ms) showing a red square move from one location to another (out of four possible locations). Next, a response screen was presented, consisting of two white squares in the same locations as the demonstrated squares. The correct response was to touch the squares in the demonstrated order. Baboons completed an average of 5545 (*SD* = 1947) Training 1 trials. To pass Training 1, subjects needed to achieve 80% accuracy within a training block, two times (non-consecutively).

For humans, Training 1 consisted of 8-trial blocks and demonstration slides were 350 ms each. This is the same training procedure that was used for baboons and the 8 trials were randomly selected from the same 24 possible square/triangle configurations. After each block, the subject's accuracy was assessed. If below 80%, the subject repeated the training level. Accuracy criteria were the same for all training levels. Humans required an average of 13.06 (*SD* = 7.68) trials to pass Training 1.

Training 2 was conducted immediately after Training 1. Demonstration slides' display times decreased to 250 ms for humans. Trials and block composition were identical to Training 1 except that four white squares were given as options during the response phase instead of two (See Fig 2.2-1 for example of four square setup). Baboons and humans completed an average of 6095 (*SD* = 2141) and 10.42 (*SD* = 8.35) Training 2 trials, respectively.

Training 3 is identical to the BASE condition in the testing phase. Demonstration slides' display times were equivalent for baboons and humans at 150 ms (Fig 2.2-1a). The experimental phase began immediately after subjects passed Training 3. Baboons and humans completed an average of 1,574 (*SD* = 1,000) and 11.70 (SD = 9.18) Training 3 trials, respectively.

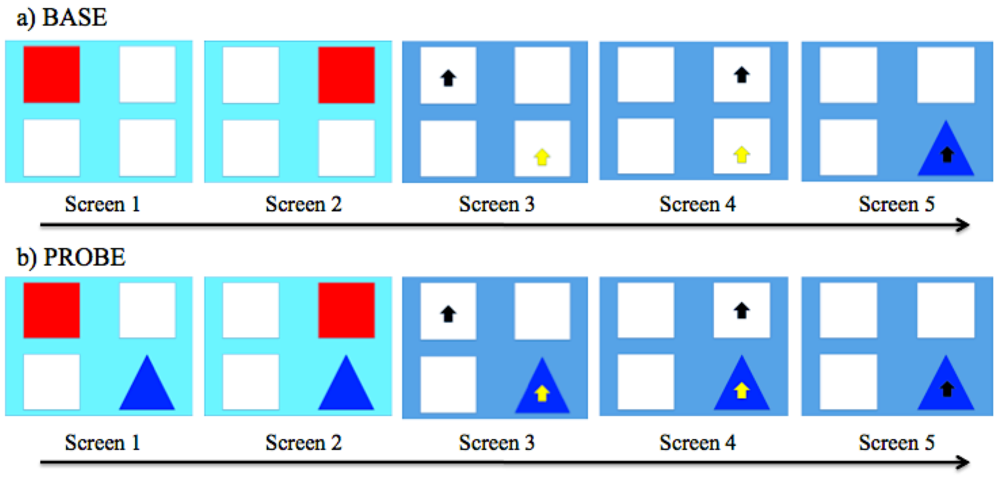


Figure 2.1 The demonstration (light blue) and response (dark blue) screens for a) BASE and b) PROBE conditions. Black arrows indicate the LS. Yellow arrows indicate where the subject could touch to use the DS. Arrows were not visible during testing. Upon touching the blue Triangle, the subject is rewarded.

2.2.4 Data Analysis

Trials were analyzed to determine whether the LS or the DS was used. Trials in which the subject sequentially touched all three response stimuli (Square1 + Square2 + Triangle) to achieve the reward were classified as having been solved by the LS. Trials in which the subject touched only two (Square1 + Triangle) or one (Triangle) response stimuli to achieve the reward were classified as having been solved by the DS. The Square1 + Triangle response was included as a DS in an attempt to conservatively maintain the dichotomous LS vs DS paradigm. For each subject, the number of trials in which the DS was used was divided by the number of correct trials completed. This yielded a DS-use ratio for both BASE and PROBE trials. Next, for each trial-type (BASE and PROBE) subjects were classified as preferring the DS (DSer) or the LS (LSer) based on this DS-use ratio. For percent DS use, the median was 0% for both BASE and PROBE conditions and the mean was 1.01% (*SD* = 2.16) for BASE and 8.01% (*SD* = 20.44) for PROBE trials. Thus, subjects who used the DS in more than 5% of BASE trials were classified as DSers in the BASE condition. Alternatively, subjects who used the DS in fewer than 5% of BASE trials were classified as LSers in the BASE condition. The same highly conservative criteria were used to classify subjects as DSers and LSers in the PROBE condition.

To look at the effects of age, humans were classified into three age groups: Children: ages 7-10 (*N* = 27, *M* = 8.44, *SD* = 1.15), Adolescents: ages 11-18 (*N* = 25, *M* = 13.52, *SD* = 2.22), and Adults: ages 19-68 (*N* = 52, *M* = 40.48, *SD*= 13.18). Age effects were not investigated in baboons, as there was extremely little response variation between subjects.

2.3 Results

2.3.1 Baboons

Mean percentages correct for BASE and PROBE were 80.7% (*SD* = 4.5) and 82.9% (*SD* = 11.0), respectively. Combined, subjects used the LS in only 3 (0.02 %) of the 1,790 PROBE trials compared to 6,898 (98.98%) of the 6,969 BASE trials. Additionally, in 20 PROBE trials

(0.11 %), the baboons touched the correct first red square but then skipped the second and proceeded to touch the blue Triangle (Note: this is later referred to as the ‘switch strategy’; See Section 3.3.2). Further, all baboons immediately switched to the DS on the first PROBE trial; the three times subjects failed to use the DS were trial numbers 22, 37, and 49. All the tested baboons therefore showed a pronounced and immediate preference for the more efficient, DS method in the PROBE condition and were classified as DSers (see Fig 2.3-1).

2.3.2 Humans

Mean percentages correct for BASE and PROBE were 91.2% (*SD* = 10.1) and 89.5% (*SD* = 11.1) respectively. Among the 104 subjects, only 21 (20.2%) used the DS in greater than 5% of PROBE trials. Of these, only 7 (6.7%) used the DS in over 50% of trials, indicating that they were able to overcome cognitive set and use the more efficient alternative method consistently. Interestingly, 50 humans (48%) used the DS at least once. Thus, even after discovering the more efficient alternative, their set was unbroken.

A Yates’ continuity corrected chi-square (used due to an expected value smaller than 5) compared the frequencies of LSers vs DSers in the two species and confirmed that the number of DSers was greater in baboons than in humans [χ^2 (1)= 35.88, $p < .001$; Figure 2.3-1]. Additionally, another Yates’ continuity corrected chi-square, indicated that there was no significant association between BASE solution strategy classification and species [χ^2 (1)= .105, $p = .746$; Figure 2.3-1].

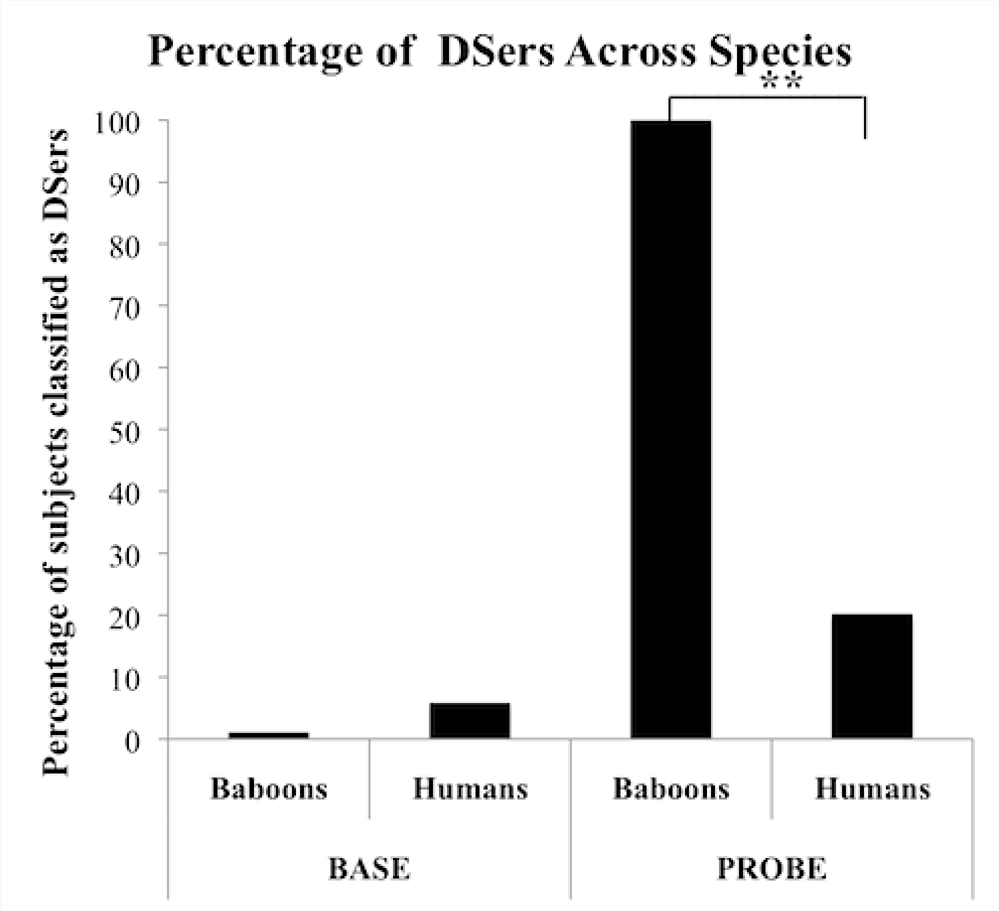


*Figure 2.2 The percentage of subjects classified as DSers in the BASE and PROBE conditions across baboons and humans. ** $p \leq .001$.*

To investigate the impact of age on DS-use, human subjects were reclassified as DSers or LSers based on their first 48 trials (24 BASE, 24 PROBE). This was done to eliminate the difference in trial number between children (who received 48 trials) and adolescents and adults (who received 96 trials). A Pearson's chi-square revealed that there was a significant association between age group and PROBE solution strategy classification in humans [χ^2 (2)= 13.32, p = .001; Fig 2.3-2]. Further, the only category in which the standardized residual was significant (2.8) was Children DSers, indicating that they were driving the effect. The association between BASE trials and age group was not significant χ^2 (2)= 1.60, p = .923 (Fig 2.3-2).

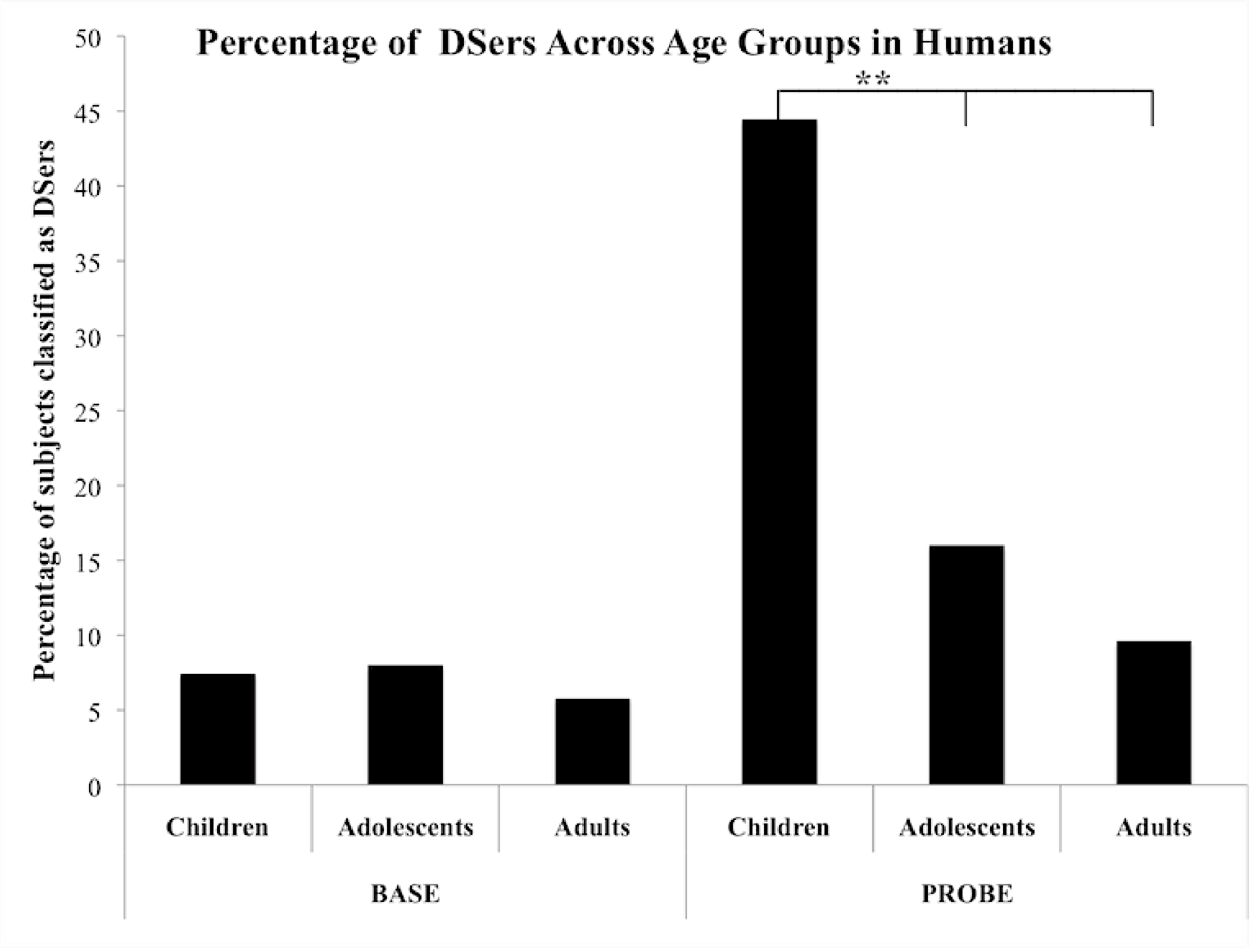


*Figure 2.3 The percentage of subjects classified as DSers in the BASE and PROBE conditions across human children, adolescents, and adults. ** $p \leq .001$*

2.4 Discussion

In this study, the first main finding was that baboons and humans responded differently on a cognitive set task. Baboons immediately broke set and adopted the more efficient DS when it became available, yet the majority of humans failed to deviate from the LS. Our second finding was that humans' ability to break cognitive set is associated with their age. Children were 3 times more likely to be classified as DSers in the PROBE condition than adolescents and 2.4 times more likely than adults. As far as we know, this is the first study to investigate cognitive set in a nonhuman species and it is one of very few to look at developmental differences in susceptibility to cognitive set in humans (Cunningham, 1965; Janzen, Maguire, & Boersma, 1976; Luchins, 1942).

Previous findings regarding age effects and cognitive set are inconsistent. Luchins (1942) found a trend (of unreported statistical significance) indicating that public school children (ages 9-14) were less able to recover from cognitive set compared to adults (ages 16-52). Cunningham (1965) tested children ages 7-12 on modified cognitive set tasks and found (minimally reported) trends indicating that older subjects were better able to overcome set. However, Janzen et al. (1976) tested children (ages 5-12) on visual set tasks and found no significant age effects. We propose that the LS-DS task is better able to compare cognitive set across ages (and species) than previous methods. While previous set tasks have involved arithmetic (Cunningham, 1965; Luchins, 1942) and alphabetic rules (Cunningham, 1965), the LS-DS required the use of a spatiotemporal rule: For the LS, identify the two demonstrated squares in their demonstrated order, then select the Triangle. This metric allowed us to a) compare set between baboons and humans and b) compare across age groups where all subjects were naïve to the task and it's rules prior to testing. Thus, the LS-DS task may be a better test of cognitive set across age groups because it does not rely on math or language skills, which are very different between children, adolescents, and adults.

Although the LS-DS task was extremely similar between humans and baboons, it was not identical. That said, differences in methodology between species should have promoted the opposite of our observed effects and thereby strengthen our findings. Humans received longer display times during training, which could have conferred increased salience to the LS. Yet, overall LS salience was heavily weighted towards baboons as they received an average of 12,915 more training trials than humans. Research suggests that increased training with a rule decreases the likelihood of participants' breaking set (Crooks & McNeil, 2009). Thus, baboons should have been less able to break set than humans based on LS experience, which was not the case.

The differential abilities of baboons and humans to break cognitive set are extreme and yet, an underlying cause is not immediately apparent. Why did the baboons immediately consider the DS, whereas humans ignored it? One hypothesis is that differences in visual and sequential processing may have conferred increased perceptual awareness of the DS to baboons. Indeed, the baboons used the DS the very first time it was available and then continued to use it in nearly every subsequent PROBE trial. To do this, they must have (a) been aware of the Triangle's premature presence in PROBE trials and (b) associated it, not the sequence as a whole, with the reward. In line with Ohshiba's (1997) findings illustrating the serial and collective search strategies of old world monkeys and humans respectively, baboons may have perceived the task's solution as a series of individual stimuli [(Square1) + (Square2) + (Triangle) = Reward] and humans may have perceived it as a collective rule [i.e., (Square1 + Square2 + Triangle) = Reward]. Thus, if baboons solved the LS-DS task with a serial search strategy, it might have allowed the DS [(Triangle) = reward] to be visually disentangled from within the LS and thus, used more effectively. Humans, on the other hand, may have used a collective search strategy and been less attentive to the Triangle's premature presence in PROBE trials. Further, Bilalić et al. (2008) found that previous experience with a solution strategy biased visual attention towards that strategy in expert chess players. However, the possible differences in visual and sequential processing of the LS-DS task between baboons and humans does not explain why only 14% of humans who used the DS at least once were able to break set. Even if it was accidental, what prevented the majority of humans from switching to the DS after discovering it?

Another explanation for humans' inability to break set is that they simply did not understand that they were allowed to. Humans' notions of how they should respond might block the use of alternative solutions. Since the classic Milgram shock experiments (1974), obedience to authority has been known to affect human behavior and this has been extended to experimenter presence and the experimental environment in general (see Rosenthal & Rosnow, 2009 for discussion). For the current study, humans' responses may have been affected by the presence of the experimenter and/or the knowledge that the task was a scientific study. It is possible that they saw the LS as the way they should solve the task based on their experience with the training and the experimental environment. Baboons, on the other hand, had free access to the testing apparatuses, without the presence of an experimenter and are likely unaffected by the experimental environment. This species difference in 'obedient' responses is supported by findings showing that following a live demonstration of how to access food from a box, humans but not chimpanzees imitated superfluous actions (Horner & Whiten, 2005). Although the current study did not measure humans' conceptual understanding of the task directly, pilot participants were asked if they had thought about touching the Triangle directly after task completion. Responses varied from "I didn't see a triangle" to "I thought it was a trap" to "Yes, and I tried it once" (See Supplementary Figure 2.2-2 in Appendix A for Pilot data). However, even the pilot-subject who tried the DS continued to use the LS afterwards, which is consistent with the 43% of non-pilot participants who "discovered" the DS yet continued to use the LS. The question now becomes: If a subject is able to see the early onset of the Triangle in PROBE trials and is willing to try touching it directly, what prevents the majority of them from adopting it as a consistent strategy?

We propose that working memory availability plays an important role in humans' persistent use of the LS. In 2007, Beilock and DeCaro found that, when under stress, humans with lower working memory availability used the direct response in Luchins' (1942) task more than humans with higher working memory. They posited that those with higher working memory were better able to remember and enact the learned rule, while those with lower working memory favored the less memory-intensive, direct response. Although the current task did not appear to induce stress in subjects, if we consider that the same working memory constraints might also have driven the increased DS preference in children, who show lower working memory skills than adults (Miles, Morgan, Milne, & Morris, 1996; Thomason et al., 2009), then our age effect becomes more coherent. The LS requires subjects to remember the locations of Square1 and Square2, while the DS only requires the subject to touch the visible Triangle. Simply stated, the LS requires working memory and the DS does not. Thus when we consider their lower working memory availability, it seems logical that more children favored the DS than adults who are presumably better equipped to handle the working memory load necessitated by the LS. This is corroborated by the comment of a 7-year-old pilot subject after he discovered the DS, "I like it when the triangle is already there because I don't have to remember the squares!" Further, baboons show overall lower working memory skills than humans (Fagot & De Lillo, 2011) suggesting that, while a serial search strategy may allow them to see the DS more readily than humans, their limited working memory could provide increased incentive to use the DS. Adults' and adolescents' persistent use of the LS may simply be a combination of inherent cognitive set and a lack of working-memory based incentive to deviate from what they've learned. This hypothesis should be explored in future studies.

In summary, the current study presents findings suggesting that baboons are less susceptible to the negative effects of cognitive set than humans. This is, as far as we know, the first comparative cognitive set study. It should be noted that in Luchins' original cognitive set task, an 'extinction problem,' where the only possible solution was the direct one, was sometimes used to enhance subjects' ability to break set. The current study did not incorporate an extinction problem but this might have an interesting effect on the observed differences. While future studies are required to more fully understand these species and age differences in ability to break cognitive set, the current study proposes that:

(1) Baboons' immediate use of the DS is facilitated by an increased ability to see the difference between the PROBE and BASE trials, which is a result of independently processing the individual components of the task sequence. Further, continued use of the DS is promoted by its minimal working memory requirements. Free from experimenter effects, baboon responses were unaffected by the experimental environment and their training with the LS.

(2) After extracting the collective LS from the training, humans' persistent use of it may have been governed by a combination of a) difficulty visually differentiating between the PROBE and BASE trials, b) consideration of how they should respond as dictated by the experimental environment and their training, and c) differences in working memory availability, with lower working memory availability promoting DS-use and higher working memory enabling LS-use.

Although nonhuman primates may encounter complex ecological, physical, or social problems, they are likely variable and not easily solvable by a single governing rule. Humans, on the other hand, are regularly faced with complex similar problems, which readily lend themselves to rule-based solutions. The adaptive benefits (or detriments) of cognitive set are not fully understood but it seems logical that set facilitates humans' ritualized problem solving. It would be interesting to address the presence of cognitive set in non-traditionally educated human populations and/or other nonhuman primate species.

3 THE EVOLUTION OF COGNITIVE FLEXIBILITY IN PRIMATES: CHIMPANZEES’ INTERMEDIATE SUSCEPTIBILITY TO COGNITIVE SET

This manuscript is in preparation for publication.

Sarah M. Pope[1,2], Joël Fagot[2,3], Adrien Meguerditchian[2,3], Michelle M. Autrey[4], William D. Hopkins[1,4]

[1] Neuroscience Institute and Language Research Center, Georgia State University, Atlanta, Georgia, USA

[2] Laboratory of Cognitive Psychology, CNRS and Aix-Marseille University, Marseille, France

[3] Brain and Language Research Institute, Marseille, France

[4] Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, Atlanta, GA, USA

Abstract

Within human problem solving, the propensity to persistently evoke a learned rule, rather than switch to a more efficient alternative is pervasive. This susceptibility to 'cognitive set' prevents optimization by biasing response patterns toward habit-based solutions. In a recent study, which utilized the nonverbal, spatiotemporal 'Learned Strategy-Direct Strategy' (LS-DS) touch screen task, baboons exhibited a striking ability to deviate from their learned strategy to utilize a more efficient shortcut. Humans, on the other hand, displayed the opposite response pattern and almost exclusively employed the familiar response. In the current study, we sought to explore the likely evolutionary trajectory of cognitive set within the primate lineage by conducting the LS-DS task with ten chimpanzees (*Pan troglodytes*). We found that, like baboons, chimpanzees used the shortcut significantly more often than humans. However, unlike either baboons or humans, we observed pronounced inter- and intra-individual variability in chimpanzees' shortcut-use. Additionally, we found that chimpanzees did not exhibit switch costs when switching between the learned strategy and the shortcut. Further, a subset of chimpanzees employed a unique solution, wherein they switched strategies mid-trial. These data indicate that chimpanzees experience an intermediate susceptibility to cognitive set on the LS-DS task. We propose that differences in abstract rule encoding may underlie differences in susceptibility to cognitive set on the LS-DS task within the primate lineage.

3.1 Introduction

Primates occupy highly dynamic environments. Decisive actions are demanded often, with outcomes that can greatly impact fitness. Adaptive behavior is predicated upon flexible strategy-use, yet such plasticity is complex. We define cognitive flexibility as the ability to incorporate both known solutions and innovated or acquired novel solutions in a contextually appropriate manner (Buttelmann & Karbach, 2017; Lehner et al., 2011). Flexible responses must integrate external environmental cues with internal inputs, such as past experience, and in the case that a past strategy is no longer the most appropriate, flexible behavior requires inhibiting that previous response and switching to a more efficient strategy. However, humans exhibit deficits in this regard.

Numerous studies describe failures by humans to deviate from a learned or familiar method in order to adopt a better alternative (Adamson, 1952; Aftanas & Koppenaal, 1962; Bilalić et al., 2008; Chrysikou & Weisberg, 2005; Crooks & McNeil, 2009; Duncker & Lees, 1945; Luchins, 1942; Luchins & Luchins, 1950; Ruscio & Amabile, 1999; Sweller et al., 1982). Yet, this propensity toward mechanized problem solving or ‘cognitive set’, is less evident in some nonhuman primate species. For instance, Pope et al. (2015) presented baboons and humans with a nonverbal ‘Learned Strategy-Direct Strategy’ (LS-DS) touch screen task. The LS-DS task begins with several sessions of training wherein subjects become proficient in using a three-step sequence (Square1→Square2→Triangle), which constitutes the learned strategy (**LS**; see Figure 1.6-1a). Once subjects consistently utilize the LS, experimental trials are presented in which subjects can use the LS (See Figure 1.6-1b) or they can use a more direct strategy (**DS** or the shortcut) by skipping Square1→Square2 and going straight for the Triangle (See Figure 1.6-1d). Pope et al. (2015) found that all 15 baboon subjects immediately switched to the DS when it became available, using it in 99.9% of trials. By comparison, only four (out of 53; 7.6%) adult humans used the DS in greater than 5% of trials.

Numerous studies aimed at identifying the origins of human cognition, including cognitive flexibility, have compared monkeys, apes, and humans (Genovesio et al., 2005; Genovesio & Wise, 2008; Manrique & Call, 2015; Rosati, 2017; Stoet & Snyder, 2008). Typically, when quantifying flexibility, studies utilize forced-switch paradigms, wherein subjects are *required* to switch strategies (Hommel & Colzato, 2017; Manrique & Call, 2015; Rosati, 2017). The most basic forced-switch tasks, discrimination reversals such as transfer index and the Wisconsin Card Sorting Task [or Dimensional Card Change Sorting (DCCS) task], reward subjects for selecting a certain response and then switch to rewarding a different response (Berg, 1948; Rumbaugh, 1971; Zelazo et al., 1996). In these tasks, monkeys and humans exhibit similar difficulties extinguishing a learned, rewarded response (Bonté, 2011; Bonte et al., 2014; Ionescu, 2012; Jordan & Morton, 2012; Manrique & Call, 2015; Piaget, 1954; Rosati, 2017; Rumbaugh, 1971; Zelazo, 2008). Yet several differences between monkeys' and humans' rule following behavior are also apparent.

Like humans, monkeys are capable of learning abstract rules (Bonté, 2011; Fagot & Bonte, 2010; Genovesio et al., 2005; Genovesio & Wise, 2008; Pope et al., 2015; Stoet & Snyder, 2008; Stoet & Snyder, 2003). However, in stark contrast to humans, monkeys reportedly do not exhibit overt trial-to-trial switch costs (Stoet & Snyder, 2003).[1] Defined as deficits in response time or accuracy associated with switching solution strategies in consecutive trials as opposed to repeating a strategy (Brass et al., 2008), switch costs are attributed to the cognitive reconfiguration allowing a new strategy to guide behavior (Lemaire et al., 2017; Luwel et al., 2009; Meiran, 1996) and are exhibited by both human children and adults (Ionescu, 2012;

Zelazo, 2008). Interestingly, these deficits were somewhat reduced if the subject knew the switch was coming (Rogers & Monsell, 1995) or had more time to prepare for it (Arrington & Logan, 2004). However, humans continued to exhibit switch costs even after 23,000 trials (Stoet & Snyder, 2007). Furthermore, although unencumbered by switch costs, monkeys did show deficits on trials which included irrelevant stimulus information, yet humans did not (Stoet & Snyder, 2003). This seems to indicate a trade-off between flexibility and distractibility, perhaps driven by species differences in the strength with which rules are encoded (Ghirlanda, Lind, & Enquist, 2017; Stoet & Snyder, 2008).

It is unclear to what extent apes are capable of flexibly adopting new strategies (Hrubesch et al., 2009; Manrique et al., 2013; Marshall-Pescini & Whiten, 2008; Price et al., 2009; Van Leeuwen et al., 2013). Apes perseverate similarly to humans in some measures (Davis, 2017; Manrique & Call, 2015; Rosati, 2017 for review); however, chimpanzees do remarkably well in inhibiting food consumption when it can be exchanged for a better reward, even after a delay (Beran, Rossettie, & Parrish, 2016). Furthermore, on simple stimulus-response discrimination reversal tasks, apes outperformed monkeys in their ability to adjust to the changing contingencies (Rumbaugh, 1971). Yet, evidence regarding their abilities to switch between more abstract rules is limited. In fact, we are aware of a single effort to elucidate abstract response flexibility in apes, in which chimpanzees (N = 6) completed a computerized Dimensional Change Card Sorting task, a simpler derivative of the Wisconsin Card Sorting Task (Ionescu, 2012; Jordan & Morton, 2012; Kirkham et al., 2003; Manrique & Call, 2015; Zelazo et al., 2003), which requires subjects to first categorize stimuli based on one stimulus dimension (e.g., shape) and then switch to using another dimension (e.g. size; Moriguchi et al., 2011). The authors reported that chimpanzees' ability to switch strategies was affected by their previous experience with the familiar strategy. However, they also noted that subjects passed the pre-switch portion of the task, on average, in only 42.9% of sessions. It is therefore unclear if even the first rule was adequately adopted, much less to what extent subjects were affected by cognitive set.

Several studies report that chimpanzees stick to a learned strategy, even after watching a demonstration of a more efficient alternative (Davis, 2017; Marshall-Pescini & Whiten, 2008). In fact, some chimpanzees were shown to persist with a learned strategy even when the alternative was easier (Hrubesch et al., 2009; Price et al., 2009). Similarly inflexible behavior has been documented in a number of baited-box experiments, in which subjects first learn one method for obtaining a reward but are then shown a better alternative (Davis, 2017; Davis et al., 2016; Gruber et al., 2011; Manrique & Call, 2015). Further, like humans, functional fixedness, wherein knowledge regarding an objects' typical function interferes with using it in an atypical manner (Adamson, 1952; Duncker & Lees, 1945), has been suggested to affect apes' problem solving tactics (Gruber, 2016; Hanus et al., 2011).

Yet, there is also compelling evidence for problem-solving flexibility in apes. In token exchange paradigms, chimpanzees flexibly switched to exchanging an unfamiliar token when it became more highly rewarded than a familiar token (Van Leeuwen et al., 2013) and (Davis, 2017) found that when an alternative solution required relatively simple modifications, chimpanzees readily switched strategies, especially when the alternative reaped a better reward. These findings are in line with others demonstrating that simple modifications to known behaviors are well within apes' capabilities (Lehner et al., 2011; Yamamoto et al., 2013).

Given the prevalence of both conservative and flexible behavior, some authors have suggested that a range of factors might influence flexibility in apes, such as how strongly the learned strategy is encoded, the relative difficulty between the familiar strategy and the alternative, and the reward difference between the two (Davis, 2017; Hopper et al., 2015; Marshall-Pescini & Whiten, 2008; Van Leeuwen et al., 2013). In contrast, each strategy on the LS-DS is performed using the exact same reach/touch motion and even stimulus locations are meaningless, as they change with each trial. Thus, the LS-DS task elucidates subjects' proclivity to switch strategies when presented with an alternative strategy that only differs from the learned strategy in efficiency (# of reaches). Thus, response-style (conservative vs flexible) on the LS-DS task is likely indicative of the cognitive representations underlying strategies and the relative ease with which they are replaced or altered.

We suggest several potential influences on the propensity to adopt the shortcut on the LS-DS task, which may differ across primate species. First, measured via hierarchical figures tasks, old world monkeys have been shown to exhibit a local perceptual bias (Deruelle & Fagot, 1998; Fagot & Deruelle, 1997; Hopkins & Washburn, 2002). In contrast, chimpanzees exhibit both global and local biases, potentially mediated by the inter-element distance (Fagot & Tomonaga, 1999; Hopkins, 1997; Hopkins & Washburn, 2002). Humans, however, exhibit a consistent global perceptual bias (Deruelle & Fagot, 1998; Fagot & Deruelle, 1997; Navon, 1977). Considering that attention to local detail might facilitate spotting the shortcut when it is available, global/local perceptual biases could affect its use.

Second, there is some evidence that old world monkeys, chimpanzees, and humans may process sequences differently (Beran et al., 2004; Biro & Matsuzawa, 1999; Fagot & De Lillo, 2011; Ohshiba, 1997). The LS-DS task requires subjects to remember the order and location of the first two items (Square1 and Square2) and then select the Triangle. Processing each step of the LS sequence as an individual component [ie. (Square1) + (Square2) + (Triangle)] or even just processing the Triangle as separable from the Squares [(Square1 + Square2) + (Triangle)], might allow subjects to more readily identify the shortcut (Triangle). On the other hand, processing the LS sequence as a collective whole [ie. (Square1 + Square2 + Triangle)], might render the DS more difficult to disentangle.

Third, although both monkeys and apes are capable of learning and utilizing abstract rules, this often requires extensive training (Beran et al., 2004; Fagot & De Lillo, 2011; Pope et al., 2015; Stoet & Snyder, 2003). Humans, by comparison, have been shown to codify rules almost immediately (Crooks & McNeil, 2009; Luchins, 1942; Luchins & Luchins, 1950) and readily apply them to novel problems (Sweller et al., 1982). Thus, the strength with which abstract rules are encoded might conceivably influence how flexibly they can be used.

The current study had two aims. First, we sought to explore chimpanzees' susceptibility to cognitive set on the LS-DS task. Second, we sought to contextualize these findings by comparing chimpanzees' response patterns to those of baboons and humans, previously reported in Pope et al. (2015). We reasoned that if chimpanzees responded similarly to humans on the LS-DS task, then limited abstract response flexibility might be considered an homologous, ancestral trait, perhaps even a byproduct of global perceptual biases or 'collective' sequential processing. However, if chimpanzees responded similarly to baboons, we might consider humans' susceptibility to cognitive set to be the outcome of a derived human characteristic, possibly a byproduct of enhanced rule encoding. Although far from a complete evolutionary framework, the current study aimed to provide the first look at a likely evolutionary trajectory of cognitive set within the primate lineage.

3.2 Methods

3.2.1 Subjects

We tested 10 captive-born chimpanzee subjects [aged 16.6–36.1 years; five females (*M* = 26.9 years, *SD* = 9.0) and five males, (*M* = 21.3 years, *SD* = 3.0)] located at the Yerkes National Primate Research Center in Atlanta, Georgia.

3.2.2 Equipment and Materials

The LS-DS was programmed using EPrime (version 2.0, Psychology Soft- ware Tools, Pittsburgh) and administered using a 19-inch touch monitor (1939L Open-Frame Touchmonitor, Elo Touch Solutions) affixed to a metal housing, which was temporarily attached to subjects' home enclosure during testing. Each testing session lasted roughly 20 minutes to avoid fatigue and subjects were rewarded with juice or small pieces of fruit or vegetables, depending on their preferences or dietary restrictions. All testing was approved by the Emory University IACUC.

3.2.3 LS-DS Training

Chimpanzees received four LS-DS Training levels. To progress to the next training level, subjects were required to achieve >80% accuracy twice. In Training 1 subjects were presented with two squares, Square1 and Square2, which flashed red (250 ms each) in sequence. To be correct, they then selected Square1 and Square2 in the demonstrated order. Accuracy was assessed after each block of 24 trials. Next, subjects progressed to Training 1.5, in which trials identical to those in Training 1 were randomly interspersed with trials in which four squares were shown and one flashed (250 ms each) red; for these, subjects then selected the square that flashed from among the four choices. Training 2 also displayed four squares, except two squares flashed (250 ms each) red in sequence and subjects were required to select Square1 and Square2 in the correct order, from among the four options. Training 3 very nearly resembled Training 2, except the demonstration squares were displayed for a shorter time (150 ms each) and after a correct Square1 → Square2 selection, a blue triangle appeared in one of the two remaining locations. For all Training 3 trials, subjects were rewarded upon touching the blue Triangle. Thus, to pass Training 3, subjects needed to be proficient in using the Square1→Sqaure2→Triangle sequence (i.e., the learned strategy or **LS**).

We encountered several difficulties during training that required procedural modifications. When the Triangle first appeared in Training 3, five of the ten chimpanzee subjects struggled to incorporate it into their response (i.e., they completely avoided it). After several sessions wherein subjects were achieving extremely low accuracy scores, we opted to briefly modify the program for these subjects, in an effort to prevent them from giving up altogether. For this modification, once a correct Square1 → Square2 selection was entered, the Triangle still appeared in one of the two remaining places but all the other squares disappeared, such that all that was left on the screen was the Triangle. Once subjects consistently touched the Triangle (*M* = 26.90; *SD* = 47.04), they were switched back to the regular version of Training 3. Additionally, in some cases, subjects' motivation appeared to dwindle substantially. When this occurred, we altered their food reward or instated a 5 second (as opposed to the normal 3 second) delay following incorrect responses. For Training 3, even if subjects were accounting for their previous selections, the likelihood of being correct in any given trial simply by chance is 1/24 (Response1 = 1/4, Response2 = 1/3, and Response3 = 1/2). Therefore, evidenced by considerably higher than chance accuracy scores (in one case a subject had achieved above 75% accuracy nine times) many subjects grasped the LS, yet failed to reach the >80% criterion twice.

After several months, we opted to adjust the Training 3 accuracy requirements such that a subject needed to achieve either greater than 80% twice, greater than 75% three times, or greater than 70% five times in order to progress to the experimental trials. Potential impacts of these alterations are discussed.

3.2.4 LS-DS Experimental Trials

Immediately after training, subjects completed 96 experimental trials, consisting of 48 PROBE and 48 BASE. In PROBE trials, the Triangle appears alongside the Square1 → Square2 demonstration and remains visible on the response screen (see Fig. 1.6-1b-d). Subjects can then either continue to use the Square1→Square2→Triangle sequence (i.e., the **LS**) *or* they can simply ignore the demonstration and select the Triangle (i.e., the **DS** or shortcut). To provide a measure of accidental DS-use within each subject, in BASE trials (which appear identical to Training 3) if subjects select the Triangle's hidden location they are scored and rewarded as if they had used the DS. Subjects randomly received each of the possible 24 configurations 4 times (2 PROBE and 2 BASE).

3.2.5 Data Analysis

We calculated a "true" measure of DS-use by subtracting each subject's BASE DS-use from PROBE DS-use, thereby accounting for within-subject error (Pope, Fagot, Meguerditchian, Washburn, & Hopkins, Submitted). For all analyses, whenever the data violated the assumptions for parametric statistical analyses, we used non-parametric analyses and reported the appropriate descriptive statistics. In accordance with our previous studies, subjects were classified as DSers if they used the DS in greater than 5% of PROBE trials (Pope et al., Submitted; Pope et al., 2015). However, we also included progressively more stringent DSer qualification criteria (greater than 25%, greater than 50%, and greater than 70%) to gain a more complete understanding of DS-use between species.

Additionally, we analyzed for switch costs associated with switching between the LS and the DS. First, we looked for deficits in response time (RT) by isolating the time between fixation response and first response (RT1) for BASE trials in which subjects repeated (BASE *stay*) or switched (BASE *switch*) their strategy and for PROBE trials in which subjects repeated (PROBE *stay*) or switched (PROBE *switch*) their strategy. All trials that were precluded by an incorrect trial and those in which the first response was incorrect were excluded. Second, we compared subjects' Response 1 accuracies between trials in which the condition repeated or was different. For example, a PROBE trial preceded by another PROBE trial was classified as PROBE *same*, a PROBE trial preceded by a BASE trial was classified as PROBE *different*, a BASE trial preceded by another BASE trial was classified as BASE *same* , and a BASE trial preceded by a PROBE trial was classified as BASE *different.* Thus, strategy choice was described as *stay*/*switch* and trial type was described as *same*/*different.*

3.3 Results

3.3.1 Training

For chimpanzees, training required an average of 295.8 days (*Min* = 27; *Max* = 465). Subject completed a *Median* of 817 (*SD* = 863.9) Training 1 trials, 2,343 (*SD* = 1,637.8) Training 1.5 trials, 8,771 (*SD* = 3,810.6) Training 2 trials, and 4,852.5 (*SD* = 6,548.0) Training 3 trials. In total, chimpanzees conducted a *Median* of 17,960 trials (*SD* = 9,989.7).

To better understand the immense number of training trials that chimpanzees needed to learn the LS, we analyzed for error patterns during Training 2. We chose to analyze Training 2 errors because, unlike Training 1, subjects selected between all four response options. Further, even by Training 3, Response 1 (*Mdn* = 69.64%) and Response 2 (*Mdn* = 66.65%) accuracies were significantly lower than Response 3 [*Mdn* = 94.05%; $\chi^2(2) = 12.60$, p = 0.002] indicating that subjects' difficulty learning the LS involved their representation of the Square1→Square2 portion of the solution, not the Triangle.

We looked at incorrect trials that were preceded by correct trials. For each Response 1, we noted whether the erroneous selection was a repetition of the previous trial, either the previous Response 1 or the previous Response 2. For example, if subjects had been rewarded in the previous trial for a Bottom Left → Top Left response, then we identified whether the erroneous response had been either a Bottom Left or Top Left selection. If driven by chance, we would expect subject's errors to correspond to these squares 33.33% of the time (1 out of the three erroneous response options). However, a one sample T-Test found that, chimpanzees' propensity to erroneously select their previous Response1 did not differ from chance (*M* = 31.05%, *SD* = 10.59). Additionally, chimpanzees erroneously re-selected their previously correct Response 2 significantly *les*s often than expected by chance [*M* = 22.49%, *SD* = 3.76; $t(9) = -6.94$, $p < .001$].

Next, we identified whether subjects' errors were, instead, due to mistakenly selecting the Square2 location, rather than the Square1 location. For example, when the demonstration depicted Bottom Left → Bottom Right and the erroneous response was Bottom Right. Indeed, chimpanzees committed this 'reversal' error-type significantly more often than expected by chance (*M* = 49.75%, *SD* = 10.10; $t(9) = 5.14$, $p = .001$).

3.3.2 DS-use

All chimpanzee (*N* = 10/10) subjects used the DS in greater than 5% of PROBE trials. In fact, 60% (*N* = 6/10) used the DS the very first time it was available and 20% (*N* = 2/10) used it every single time it was available. Additionally, chimpanzees employed another, previously unreported strategy. In addition to the LS (Square1 + Square2 + Triangle; Figure 1.6-1b) and DS (Triangle; Figure 1.6-1d), the task permits a third strategy wherein subjects seemingly initiate the LS by selecting the first square but then switch to the DS by skipping the second square and instead selecting the Triangle (Square1 + Triangle; Figure 1.6-1c). Because it occurred so infrequently, this 'switch strategy' (**SS**) was not analyzed separately from the DS in the previous studies with baboons and humans (Pope et al., 2015). However, the current study found that chimpanzees used the SS in an average of 12.4% (*Max* = 35.7%) of trials.

3.3.3 Impact of Strategy on Performance

To assess the impact of strategy on performance, we investigated average accuracies in BASE and PROBE trials. Recall that LS, SS, and DS responses are each considered correct, only varying on their relative efficiencies. PROBE (*Mdn* = 95.83, *SD* = 10.22) accuracy was significantly higher than BASE (*Mdn* = 77.08, *SD* = 5.01) accuracy (*Z*= -2.501 , *p* = .012).

3.3.4 Switch Cost Analysis

A repeated measures ANOVA determined that there were no effects of either condition (BASE/PROBE) or strategy type (*stay/switch*) on subjects' RT1s. Additionally, a related samples Friedman's test showed no effect of trial type (*same /different*) on Response 1 accuracy; however, there was a significant effect of condition [$\chi^2(3, N = 10) = 12.59, p = 0.006$].

BASE *same* trials were significantly less accurate than PROBE *different* and PROBE *same*, although BASE *different* trials did not differ from any trial type. Thus, chimpanzees showed no switch costs in either RT or accuracy; however, as described above, they did exhibit improved accuracy on PROBE compared to BASE trials.

3.4 Comparative Analyses

Here we aimed to compare chimpanzees' responses on the LS-DS task to those of baboons and humans, which were previously collected and reported in Pope et al. (2015).

3.4.1 Subjects and General Procedure

Baboon data were collected from 15 socially housed baboons [aged 1.8–9.3 years; six males (*M* = 5.3 years, *SD* = 2.68) and nine females (*M* = 5.1 years, *SD* = 2.36)] located at the CNRS Station de Primatologie in Rousset-sur-Arc, France. The task was administered via ten automated learning devices for monkeys (ALDMs; Fagot & Paleressompoulle, 2009) and was approved by the local ‘‘Provence Alpes Cote d'Azur'' ethics committee for experimental animal research. In the previous report, children conducted only 48 trials (Pope et al., 2015).[2] Therefore we focused on the adult human data, which were collected from 53 participants (aged 18–68 years; 23 males (*M* = 39.0 years, *SD* = 13.7) and 30 females (*M* = 40.8 years, *SD* = 13.4) at a temporary testing booth set up at Zoo Atlanta in Atlanta, Georgia, USA. Human testing was approved by the Zoo Atlanta Research Committee and the Georgia State University Institutional Review Board.

3.4.2 Analysis

For all species, analyses were performed on subjects' first 96 trials, comprised of 48 PROBE and 48 BASE trials. For our original assessment of cognitive set between baboons and humans (Pope et al., 2015), we calculated DS-use based on the percentage of trials in which subjects used the DS in PROBE trials and compared it to BASE DS-use. However, for the current analyses we used the same 'true' measure of DS-use described in Section 3.2.5.

3.5 Comparative Results

3.5.1 General Species Differences

Average accuracy, as shown by a Kruskal-Wallis test, significantly differed between chimpanzees ($M = 83.8$, $SD = 6.3$) and humans [$M = 90.7$, $SD = 9.9$; $H(2) = 11.89$, $p = 0.000$] but not between baboons ($M = 88.3$, $SD = 8.8$) and humans or baboons and chimpanzees. Additionally, Mann-Whitney tests determined that sex did not influence DS-use for any species (baboons: $U = 25.5$, $p = .864$; chimpanzees: $U = 14.0$, $p = .841$; humans: $U = 404.5$, $p = .213$).

3.5.2 Species Differences in Training

Kruskal-Wallis tests confirmed that across all training levels, baboons and chimpanzees completed significantly more trials than humans: Training 1 [$H(2) = 53.84$, $p < 0.001$], Training 2 [$H(2) = 70.63$, $p < 0.001$], Training 3 [$\mathrm{H}(2) = 59.80$, $p < 0.001$], and the total number of training trials [$H(2) = 52.74$, $p < 0.001$; Table 3.5-1]. Note that chimpanzees completed an extra training level, Training 1.5, which is included in the total number of training trials. Combining Training 1.5 with Training 1 did not alter the above finding ($H(2) = 53.84$, $p < 0.001$). These differences should be interpreted cautiously given that the pass criteria differed for each species.[3] We found no significant correlations between the number of training trials (for any level) and DS-use for any species: Training 1: chimpanzees ($r_s(8) = .231$, $p = .522$), baboons ($r_s(13) = .062$, $p = .829$), humans ($r_s(51) = -.108$, $p = .439$); Training 2: chimpanzees ($r_s(8) = .407$, $p = .243$), baboons ($r_s(13) = .313$, $p = .625$), humans ($r_s(51) = .151$, $p = .281$); Training 3: chimpanzees ($r_s(8) = .164$, $p = .651$), baboons ($r_s(13) = .224$, $p = .422$), humans ($r_s(51) = -.051$, $p = .713$); Total Training chimpanzees ($r_s(8) = .438$, $p = .205$), baboons ($r_s(13) = .095$, $p = .739$), humans ($r_s(51) = -.129$, $p = .357$).

Table 3.1 Median number of training trials for Training 1, Training 1.5, Training 2, Training 3 and Total Training for all species. SD in parentheses.

	Training 1	Training 1.5	Training 2	Training 3	Total Training
Baboons	6,084	----------	6,108	1,345	14,115
	(1,947.1)	----------	(2,309.1)	(999.5)	(4,345.7)
Chimps	817	2,343	8,771	4,852.5	17,960
	(863.9)	(1,637.8)	(3,810.6)	(6,548.0)	(9,989.7)
Humans	16	----------	8	8	32
	(8.6)	----------	(10.5)	(7.7)	(20.2)

3.5.3 Error Analyses

Next, like in Section 3.3.1, we analyzed for error patterns during Training 2. Recall that we focused on incorrect trials that were preceded by correct trials. However, only 6 humans (none of which were DSers) committed eligible errors ($M = 6.66$) and thus, we did not analyze humans' training errors further. One sample T-Tests found that baboons erroneously reselected their previously correct Response 1 ($M = 23.55\%$, $SD = 7.76$; $t(14) = -4.882$) and Response 2 ($M = 14.56\%$, $SD = 3.45$; $t(9) = -21.06$, $p < .001$) significantly *les*s often than expected by chance.

Further, similarly to chimpanzees, baboons committed the ‘reversal’ error-type significantly more often than expected by chance (M = 49.29%, SD = 12.60; $t(9) = 4.91$, $p < .001$).

3.5.4 Cross-species Differences in Shortcut-Use

A Pearson’s chi square confirmed that the proportion of baboons (100%) and chimpanzees (100%) that were classified as DSers differed significantly from humans (7.5%), but not from each other [$\chi^2(2, N = 78) = 62.17$, $p < .001$; Table 3.5-2a]. Additionally, the proportions of chimpanzees and baboons that used the DS in more than 25% [$\chi^2(2, N = 78) = 58.02$, $p < .001$] and more than 50% ($\chi^2(2, N = 78) = 51.91$, $p < .001$; Table 3.5-2a) of trials were significantly higher than humans, but again did not significantly differ from each other. In the 75% DSer classification, all species significantly differed from one another [$\chi^2(2, N = 78) = 51.09$, $p < .001$; Table 3.5-2a].

The number of subjects who used the DS the very first time it was available differed significantly between all three species: baboons had the highest proportion (100%), followed by chimpanzees (60%), and then humans (5.7%) [Fisher’s Exact: $\chi^2(2, N = 78) = 55.61$, $p < .001$]. Further, the proportion of baboon (53.3%) and chimpanzee (20%) subjects who used the DS every time it was available significantly differed from humans (0%), but not from each another [Fisher’s Exact: $\chi^2(2, N = 78) = 26.65$, $p < .001$]

The SS was not distinguished from the DS previously (Pope et al., 2015), because humans and baboons used it in less than 1% of trials on average (baboons: M = 0.80%; Max = 6.25%; humans: M = 0.49%; Max = 9.52%). A Pearson’s chi square confirmed that the proportion of chimpanzees classified as SSers (50.0%) was significantly larger than the proportion of baboons (6.7%) and humans (1.9%), which did not differ from each other [Fisher’s

Exact: $\chi^2(2, N = 78) = 15.68, p < .001$; Table 3.5-2b). Further, the proportion of chimpanzees (30%) who used the SS in more than 25% of trials was significantly higher than humans (0%) but baboons (0%) did not differ significantly from either other species [Fisher's Exact: $\chi^2(2, N = 78) = 11.20, p = .002$; Table 3.5-2b]. No chimpanzees used the SS in greater than 50% of trials.

Table 3.2 Proportion of each group which used the a) DS or the b) SS in greater than 5%, 25%, 50%, and 75% of trials. Superscripts identify statistically distinguishable groupings.

a) **DS**

	>**5%**	>**25%**	>**50%**	>**75%**
Baboons	100^{a}	100^{a}	100^{a}	93.3^{a}
Chimpanzees	100^{a}	90.0^{a}	70.0^{a}	30.0^{b}
Humans	7.5^{b}	7.5^{b}	7.5^{b}	3.8^{c}

b) **SS**

Baboons	6.7^{a}	$0^{a,b}$
Chimpanzees	50.0^{b}	30.0^{b}
Humans	1.9^{a}	0^{a}

3.5.5 Impact of Strategy on Performance

To assess the impact of strategy on performance, we investigated subjects' average accuracies in BASE and PROBE trials. Wilcoxon Signed-Ranks tests showed that, like chimpanzees (and described in Section 3.3.3), baboons' PROBE ($Mdn = 97.08$, $SD = 3.50$) accuracies were significantly higher than their BASE ($Mdn = 79.58$, $SD = 15.87$) accuracies ($Z = -3.411$, $p = .001$). However, for humans, PROBE ($Mdn = 91.67$, $SD = 10.33$) accuracy was significantly lower than BASE ($Mdn = 93.75$, $SD = 10.03$) accuracy ($Z = -2.151$, $p = .031$), indicating that although they typically do not use the DS, the presence of the Triangle in PROBE trials may be distracting for humans.

3.5.6 Switch Cost Analysis

We also analyzed baboons' switch costs, as they were not assessed in the previous report (Pope et al., 2015). For baboons, a related samples Friedman's test showed no effect of trial type (*same/different*) on Response 1 accuracy. Although, similarly to chimpanzees, there was a significant effect of condition [$\chi^2(3, N = 15) = 30.15, p < 0.001$], which (as previously described) is indicative of their enhanced accuracy on PROBE compared to BASE trials (Table 3.5-3). Unfortunately due to experimenter error, baboon response times were not accurately recorded and could not be assessed for switch costs.

For the humans, only 4 subjects reliably used the DS (i.e., switched strategies); thus, we were unable to statistically assess for switch costs in humans. However, we report a descriptive account of the 4 human subjects' RT and accuracy switch costs (Table 3.5-3). Note: there was a significant difference between chimpanzees' and humans' BASE RT1s [Chimp: $M = 315.47$, $SD = 81.34$; Human: $M = 622.24$, $SD = 164.47$; $F(1,61) = 32.94$, $p < .001$)], RT2s [Chimp: $M = 779.14$, $SD = 148.32$; Human: $M = 227.19$, $SD = 77.83$; $F(1,61) = 304.8$, $p < .001$)], RT3s [Chimp: $M = 846.59$, $SD = 88.21$; Human: $M = 509.75$, $SD = 71.34$; $F(1,61) = 172.5$, $p < .001$)].

Table 3.3 For BASE/PROBE conditions, mean response time (ms) and median accuracy (%) for subjects' Response 1 in stay/switch strategy types and same/different trial types. SD in parentheses.

Response Time

	BASE *stay*	BASE *switch*	PROBE *stay*	PROBE *switch*
Baboons	- -	- -	- -	- -
Chimps	305.1 (79.6)	317.8 (153.6)	322.2 (85.9)	329.6 (101.9)
Humans	485.2 (88.6)	487.0 (111.9)	443.8 (128.2)	508.9 (189.5)

Accuracy

	BASE *same*	BASE *different*	PROBE *same*	*PROBE different*
Baboons	95.8 (2.8)	94.7 (3.9)	100 (1.8)	100 (1.1)
Chimps	87.2 (8.3)	90.0 (6.0)	95.5 (4.8)	94.9 (6.7)
Humans	97.8 (4.1)	94.6 (5.2)	95.3 (2.4)	96.3 (3.9)

3.6 Discussion

The interplay between flexible strategy updating and the maintenance of existing solution strategies presents an interesting opposition. On one hand, if an existing representation is too concrete and impervious to contextual inputs, then newly relevant information is ignored, resulting in a familiar but potentially inefficient response. On the other hand, if the representation is too easily perforated, then response efficiency is impaired via distractability by irrelevant or even maladaptive information (Cools, 2008; Hommel & Colzato, 2017; Roberts, 2008). For example, if a foraging animal is unable to flexibly adapt to shifting environments, they might starve; yet if foraging habits are too flexible, a novel and potentially poisonous food item might be consumed (Rosati, 2017). Optimal behavior requires a balance between persistent and flexible behavior and should be task-dependent (Brosnan & Hopper, 2014; Chrysikou et al., 2013).

3.6.1 Implications for Cognitive Flexibility in Chimpanzees

In this study we found that chimpanzees exhibited an intermediate susceptibility to cognitive set. On one hand, their shortcut-use resembled baboons,' in that it was far greater than humans'. However, as evidenced by their enhanced inter-individual variation in strategy selection (Figure 3.6-1) as well as their use of the SS, chimpanzees are still influenced by habit-based responses to some extent. What might facilitate this semi-flexible response-style and the implications for the evolution of cognitive set are now discussed.

As we described earlier, previous studies report considerable conservatism in chimpanzees. Specifically, compared to naïve individuals, subjects trained to use a certain solution are less able to use a better alternative once it is available (Davis, 2017; Hrubesch et al., 2009; Manrique & Call, 2015; Marshall-Pescini & Whiten, 2008; Price et al., 2009). However, there are several (often necessary) pitfalls of non-computerized tasks, which might suppress flexibility.

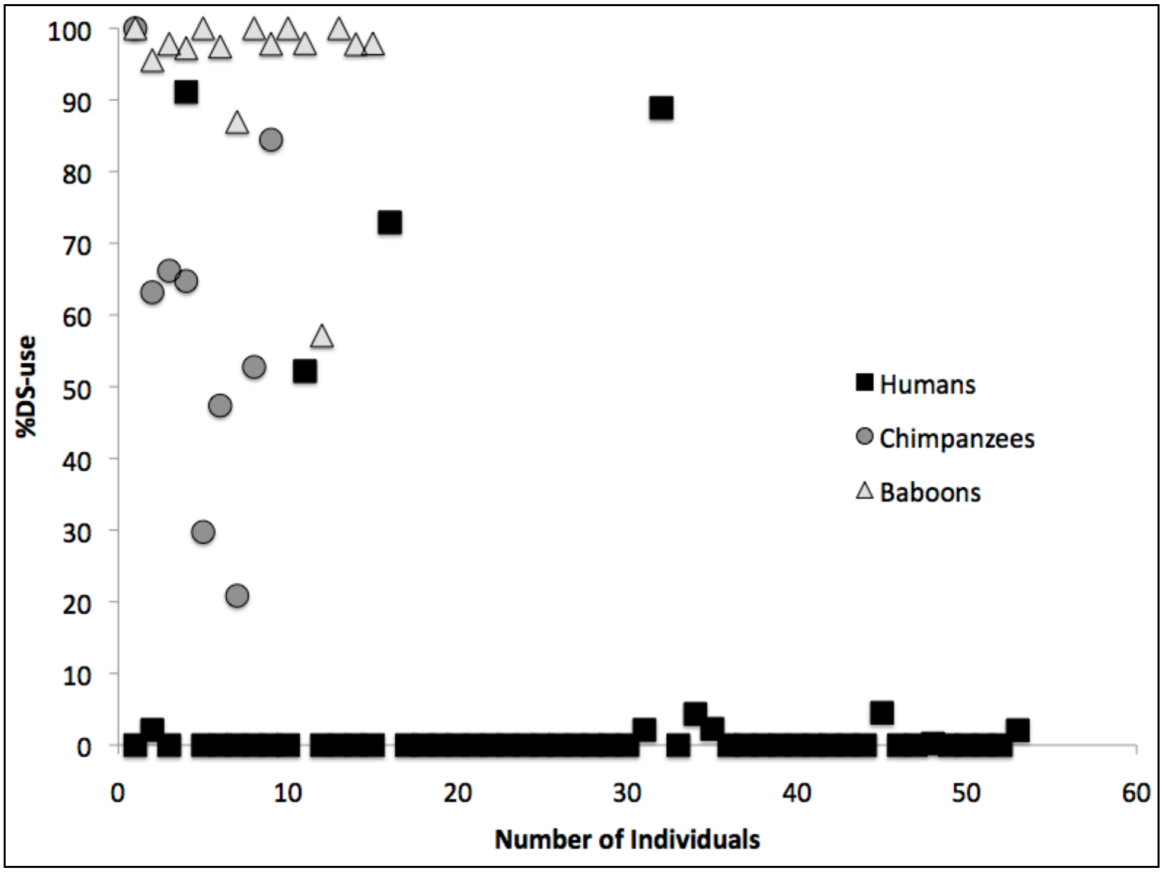


Figure 3.1 The proportion of each subjects' PROBE trials in which the DS was used, accounting for within subject error.

First, a portion of the social learning studies depicting conservatism in apes may have inadvertently negatively reinforced the alternative strategy during training by 'locking' it or rendering it ineffective (Davis et al., 2016; Manrique & Call, 2015; Price et al., 2009). This is important because if subjects attempt to use the alternative during training (the reason it is locked) and are met with failure, subsequent avoidance of the alternative solution could just be a product of their own personally unrewarding experience with that strategy (i.e., it did not work).

Second, many of these studies rely on a model to demonstrate the more efficient alternative (Davis, 2017; Dean et al., 2012; Gruber et al., 2011; Hanus et al., 2011; Hrubesch et al., 2009; Marshall-Pescini & Whiten, 2008; Price et al., 2009), introducing the possibility that conservatism may not be derived from cognitive inflexibility but from a failure to extract the relevant information from (or even attend to) the demonstration. In the LS-DS task, subjects are not shown how to enact the shortcut. DS-use is entirely self-motivated therefore our findings are not confounded by social learning ability.

Finally, many of the tasks which describe conservative behavior in chimpanzees state that the behavior is clearly within the species' repertoire (Marshall-Pescini & Whiten, 2008) and thus, would be readily utilized if subjects were not influenced by a more familiar solution. However, what may be present in the species' repertoire is not necessarily present within the individual's and inexperience with specific manipulations, especially those involving fine motor control or tool-use, should not be discounted (Dean et al., 2012; Gruber et al., 2011). The time spent learning how to perform the seemingly more efficient alternative, may easily render it less so. By comparison, the LS, DS, and SS strategies enlist motorically identical actions. Our findings clearly demonstrate that, under certain conditions, chimpanzees are capable of flexibly switching between abstract response strategies. However, the extent to which this carries over into more naturalistic settings should be elucidated in future endeavors.

3.6.2 Species Differences in Cognitive Flexibility on the LS-DS Task

By design, the LS-DS task is meant to promote shortcut use in PROBE trials. Accordingly, the DS is more desirable than the LS in several ways. First, the DS does not rely on working memory. Once the Triangle appears it remains on the screen until subjects select a response. In PROBE trials, after the demonstration is over, all that remains on the screen are three blank response Squares and the Triangle. Thus, the DS is more salient *and* less cognitively demanding than the LS. Furthermore, the Triangle is highly associated with reward. Every time it is selected, throughout Training 3 and the experimental trials (both BASE/PROBE), subjects receive a reward. This is in stark contrast to the squares, which are only directly associated with reward before Training 3 and often, if erroneously selected, lead to the incorrect screen, a timeout, and no reward. Indeed, by employing the SS, chimpanzees are effectively overriding an initial – perhaps habit-based – LS response in order to select the Triangle. In fact, subjects (of all species) which used the shortcut the very first time it was available (baboons *N* = 15/15, chimpanzees *N* = 6/10 , humans *N* = 3/53) clearly recognized the Triangle's role as a direct path to reward.[4]

Given the task design, it is not surprising when subjects, of any species, prefer the DS. What is remarkable is humans' robust proclivity for LS-use. Although chimpanzees' global perceptual biases or 'collective' sequential processing could have contributed to SS-use or reduced DS-use, they far from prevented chimpanzees from using the shortcut. In fact, overall, chimpanzees responded most similarly to baboons. We will next explore why humans' but not baboons nor (to a large extent) chimpanzees, are affected by cognitive set in the LS-DS task.

Thousands more training trials were needed for chimpanzees (*Min* = 2,784; *Max* = 36,966) and baboons (*Min* = 5,043; *Max* = 20,060) to learn the LS, compared with humans (*Min* = 24; *Max* = 152). We suggest that LS-use may be aided by humans' enhanced ability to encode strategies either through verbal encoding (Ghirlanda et al., 2017), heightened working memory (Fagot & De Lillo, 2011; but see: Inoue & Matsuzawa, 2007), or chunking strategies (Gobet et al., 2001; Kolodny et al., 2015). In fact, many of the human subjects were able to learn the rule after only 8 trials – a quarter of the total possible configurations – illustrating that, once learned, humans are capable of accurately applying the LS to novel situations (Pope et al., Submitted; Pope et al., 2015). On the other hand, chimpanzees and baboons required many encounters with the same trial configurations before they began to accurately respond. Chimpanzees' and baboons' errors seem to have been driven, to a large extent, by erroneously reversing the order of the demonstrated squares (i.e., selecting Square2 first). This reversal error is in line with recent findings which suggest that humans exhibit a pronounced advantage when it comes to sequential encoding (Ghirlanda et al., 2017). Indeed, a less firmly encoded LS might reasonably facilitate baboons' and chimpanzees' enhanced ability to replace it.

In line with this assertion, the current study found that chimpanzees, unlike humans (Pope et al., Submitted), did not exhibit switch costs on the LS-DS task. As far as we are aware, this is the first study to investigate deficits associated with switching between abstract rules in chimpanzees. Stoet and Snyder (2003) posited that switch costs are a product of firmly encoded rules, such that greater cognitive effort is required to suppress the previous strategy and activate the current. Additionally, less entrenched rules have been hypothesized to amplify distractibility (Stoet & Snyder, 2008), which for the LS-DS task might further promote the use of the shortcut.

That being said, chimpanzees, but not baboons, utilized the SS, suggesting that chimpanzees were influenced by the habit of LS-responding to a greater extent than baboons. We posit that differences in how old world monkeys and apes group the components of the LS could have affected their relative abilities to separate the shortcut from the sequence as a whole. Although somewhat unclear, there is some reason to think that baboons might identify the individual components of a sequence, while chimpanzees might process the sequence as a whole (Ohshiba, 1997); however, this is up for some debate (Beran et al., 2004). Similarly, in a matchto-sample task, baboons seemed to encode only the relevant aspects of a stimulus; yet, humans encoded irrelevant dimensions as well (Fagot, Kruschke, Depy, & Vauclair, 1998). Furthermore, in a serial recall task, baboons but not humans benefitted from sequence structure (Fagot & De Lillo, 2011). Thus, it seems reasonable to suspect that chimpanzees may exhibit intermediate susceptibility to cognitive set on the LS-DS task because, unlike humans, they are not verbally encoding their solution strategy but also differ from baboons in the way the LS is encoded. Future efforts aimed at disentangling these influences might compare cognitive set between monkeys and apes utilizing a non-sequential task.

3.6.3 Evolutionary Implications

Distinct response patterns are observed between baboons, chimpanzees, and human adults on the LS-DS task. We suggest that these may be explained by a trade-off between how quickly abstract rules can be learned and the degree to which they can be flexibly used, a process potentially governed by differences in strategy encoding.

In summary, the current study found that, like baboons, chimpanzees used the shortcut significantly more often than humans. However, they did not exhibit switch costs when shifting between the LS and the DS. Furthermore, they sometimes employed a unique solution, wherein they switched strategies mid-trial (the SS). These data suggest that chimpanzees' experience an intermediate susceptibility to cognitive set on the LS-DS task.

Footnotes

[1] Note that Huguet, Barbet, Belletier, Monteil, and Fagot (2014) *did* observe RT costs when baboons switched between trials which did or did not conflict with their learned strategy; however, this was only observed for the longest RTs (bin 5/5) and only when tested in isolation. Perhaps indicating that subjects were more distracted by the shift between the trial types. If these costs were associated with a shift in cognitive approach, we would expect them to appear within immediate responses, rather than only the slowest response times.

[2] In our original assessment of cognitive set between baboons and humans (Pope et al., 2015), we calculated DS-use based on the percentage of trials in which subjects used the DS in PROBE trials and compared it to BASE DS-use. In subsequent analyses (including the current report and (Pope et al., Submitted)), we opted to calculate a "true" measure of DS-use by subtracting subjects' BASE DS-use from PROBE DS-use. Additionally, our original analyses did not distinguish between SS-use and DS-use, they were both classified as DS responses. When we separate these strategies and re-run our analyses on the first 48 trials for each subject, the main finding, that baboons used the shortcut significantly more than humans, is preserved [$\chi^2(1, N = 104) = 49.63$ $p < .001$]. However, the age-related findings suggesting that children (ages 7-10; n = 27; DSers = 18.5%) used the DS more than adolescents (ages 11-18; n = 25; DSers = 9.6%) or adults (ages 19-68; n = 52; DSers = 4.0%) were no longer significant [$\chi^2(2, N = 104) = 2.995$ $p = 0.224$; note: two cells had expected counts less than 5]. Further, children's SS-use (25.9%) was not significantly different from adolescents' (12.0%) or adults' (7.7%). This was also true if we combined DS and SS use into a single measure, indicating that erroneous DS-use may have driven the originally reported statistical difference in children.

[3] Baboons were required to achieve 80% or more, twice, assessed after each block of 96 trials. Humans were required to achieve 80% or more, assessed after each block of 8 trials. Chimpanzees were required to achieve 80% or more, twice, assessed after each block of 24 trials. However, because we altered the pass criteria partially through chimpanzees' training, some of the subjects progressed immediately after the decision was made and therefore had already completed many more training trials than some of their peers were required to.

[4] Yet, notably, half of the chimpanzee subjects avoided selecting the Triangle when it first appeared in Training 3 (one subject even ran away from the touch screen the very first time it appeared), only selecting it when a modified version of the task removed all other options. This suggests that chimpanzees' flexibility on the LS-DS task was reliant upon their familiarity with the Triangle.

4 BREAKING COGNITIVE SET: SEEING AN ALTERNATIVE STRATEGY IS NOT THE SAME AS LOOKING FOR ONE

This manuscript is in preparation for publication.

Sarah M. Pope[1,2], David A. Washburn[3], William D. Hopkins[1,4]

[1] Neuroscience Institute and Language Research Center, Georgia State University, Atlanta, Georgia, USA

[2] Laboratory of Cognitive Psychology, CNRS and Aix-Marseille University, Marseille, France

[3] Language Research Center, Department of Psychology, Georgia State University, Atlanta, GA

[4] Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, Atlanta, GA, USA

Abstract

By applying learned rules, humans are able to solve problems with minimal cognitive effort. Yet this sort of habit-based problem solving may readily foster cognitive inflexibility or set. This occurs when an alternative – even more efficient – strategy is masked by a known, familiar solution. Here we sought to explore how cognitive set occludes alternatives, by identifying whether better solutions are a) visually overlooked or b) seen but disregarded. Subjects' eye movements were tracked while they conducted the computerized LS-DS task, which measures an individual's ability to depart from a learned strategy (LS) in order to adopt a more efficient, direct strategy (DS or the shortcut). Indeed, we found that subjects fixated on the DS prior to responding, indicating that they did not simply overlook the alternative strategy; yet, the vast majority (86.2%) did not adopt it, instead continuing to use the LS. Next, subjects were shown a video demonstrating either the DS (Informed) or the familiar LS (Control). In subsequent trials, Informed subjects fixated more on the DS prior to responding, were more likely to use the DS after seeing it, and were more likely to be classified as shortcut-users (69.0%) than Controls (6.9%). Thus, unsurprisingly, after learning that a more efficient alternative was possible, many subjects were able to break away from their familiar solution. That being said, 31.0% of subjects continued to use the learned strategy despite viewing a demonstration of the shortcut. We discuss implications for rule-based problem solving and the potential impacts on cognitive flexibility.

4.1 Introduction

Codified solutions enable humans to solve complex problems with minimal cognitive effort. However, once learned, rules may be blindly applied with little consideration for alternative strategies. In 1945, Max Wertheimer noted that children who had recently learned a method of finding the area of a parallelogram were unable to solve an identical problem after it had been rotated 90°. Subjects either did not try the new problem, stating that they had not learned how to solve it, or continued to apply their learned method, albeit unsuccessfully (Wertheimer, 1945). In fact, humans of all ages have exhibited problem-solving deficits stemming from over-reliance on familiar strategies or learned rules.

In another classic example, subjects are presented with a set of arithmetic 'water jar' problems all solvable by the same, four-step solution; however, the 7th problem can be solved not only by the learned solution but also by a more efficient alternative (Luchins, 1942). Luchins (1942; 1950) tested thousands of subjects, of all ages and under various manipulations, and found that consistently the majority of subjects did not use the more efficient alternative. Termed 'cognitive set,' the propensity, for known solutions to occlude alternative - even more efficient – strategies, has been demonstrated across a wide variety of tasks (Bilalić et al., 2008; Chrysikou & Weisberg, 2005; Lemaire & Leclere, 2014; Luchins, 1942; Luchins & Luchins, 1950; Pope et al., Submitted; Pope et al., 2015; Sweller et al., 1982), yet little is known regarding the underlying causes.

One hypothesis is that cognitive set biases visual search during problem solving, such that only stimuli relevant to the familiar method are perceived. In other words, once a strategy is adopted, alternatives are simply overlooked. For example, Bilalić et al. (2008) found that expert chess players, after finding one possible move, reported that they were looking for another, yet subjects' gaze indicated that were continuing to look at the solution they had already found. In Luchins' task, efforts to increase the saliency of the alternative strategy were met with some success (Luchins, 1942; Luchins & Luchins, 1950; Sweller & Gee, 1978), yet to what extent these findings were conflated with other changes, like increased difficulty of the learned strategy, are unclear.

Another hypothesis is that cognitive set may arise from an inability to conceptually identify the alternative as a viable solution strategy. Simply, subjects may not realize that they can or should use it. Knoblich et al. (2001) presented subjects with 'matchstick' arithmetic problems, in which they were required to alter the problem space (by moving one matchstick) to balance equations. Subjects struggled with problems requiring them to disassemble meaningful components (e.g., using Roman numeral X to create V) or alter operators (e.g., using + to create =), which the authors suggested may stem from their prior experience with mathematics. Another example comes from the Luchins' water jar task. Luchins (1942) found that instructing subjects to write "Don't be blind" increased their use of the more efficient alternative. This was also true when subjects were given an extinction problem that could *only* be solved by the alternative strategy. However, a shocking number of subjects proclaimed that the extinction problem was unsolvable (note: this problem was easily solved by naïve individuals), demonstrating the remarkable constraints imposed by the learned rule on adaptive problem solving.

We recently reported similar results using the computerized LS-DS task, a nonverbal nonmathematical adaptation of Luchins' task, which first trains subjects to utilize a three-step sequence (Square1→Square2→Triangle) and then presents them with an opportunity to use a more efficient shortcut (Triangle; Pope et al., 2015). Susceptibility to cognitive set is measured by subjects' ability to depart from their learned strategy (**LS**) in order to adopt the more direct strategy (**DS** or the shortcut). Interestingly, baboons and chimpanzees seemed relatively unaffected by cognitive set on the LS-DS task, yet less than 10% of American adults were able to employ the shortcut in more than 5% of trials (Pope et al., 2015).

Here, we sought to identify how perceptual and conceptual constraints impact susceptibility to cognitive set. By tracking subjects' gaze while they completed the LS-DS task we sought to identify if and to what extent persistent LS-use may be driven by perceptual constraints. In other words, do subjects see the shortcut and just not use it or, do they not even see it? Further, we measured subjects' shortcut-use following a video demonstration of either the LS or the DS to investigate how subjects' conceptual understanding of the DS as a potential solution might influence their ability to use it. Rule-based strategies provide a framework from which many problems may be solved; however rules are only truly useful when flexibly applied. The current study aimed to elucidate the factors contributing to cognitive set, in an effort to promote efficient selection between learned rules and alternative strategies that co-occur within a problem space.

4.2 Methods

4.2.1 Subjects

Data were collected from 58 subjects (52 females and 6 males) recruited from the pool of undergraduate students at Georgia State University by posting the study on the SONA Experiment Managements System. Subjects were tested on the Georgia State University campus in a room with dimmed lights. All subjects were above 18 years of age (*Max* = 37, *M* = 20.47, *SD* = 4.23). Fourteen subjects were not included in the analyses as a result of either technical malfunctions (*N* = 12), their accuracy being below 65% during experimental trials (*N* = 1; 62.5%), or an inability to pass the training trials (*N* = 1).

4.2.2 General Methods

Subjects were seated approximately 60cm from a 19inch monitor (1280 x 1040 Native Resolution; 1915L Desktop Touchmonitor, Elo Touch Solutions). Using the Eye Tribe Tracker (The Eye Tribe) subjects completed a 16-point gaze calibration. Next, OpenSesame (version 3.1.1; OpenSesame Experiment Builder; Mathot, Schreij, & Theeuwes, 2012), was launched and using the PyGaze plugin (version 0.6.0a16; default settings), subjects' gaze was calibrated again. Next, the feedback screens (and accompanying sounds) indicating correct and incorrect responses were demonstrated and subjects were told that they would need to 'select the shapes to figure out the right answer.' Following this, using the PyGaze drift-correct feature, subjects were instructed to look at the fixation cross while pressing the SPACE bar to start each trial. At the beginning of each trial, if subjects' gaze was not detected as being directed at the fixation cross (within a 1.5° threshold), an error sound was elicited and subjects remained on the fixation screen until it was. This ensured that subjects' gaze started from the same location in every trial and that detection remained accurate throughout the testing session. If, at any point, subjects struggled to move past the fixation screen, the experiment was paused and gaze was recalibrated. No further instructions were provided and the experimenter remained in an adjacent room (out of sight) unless recalibration was required.

4.2.3 The LS-DS Task

A full description of the LS-DS task can be found in Pope et al. (2015); however, briefly: the task is comprised of three training levels followed by 96 experimental trials. In Training 1, two squares flash red (250ms each) in sequence and subjects must reproduce this demonstration by selecting them in the correct order. In Training 2, two out of four squares flash red (200ms each) in sequence and subjects must reproduce this demonstration by selecting the two correct squares (now out of four options), in the correct order. In Training 3, two out of four squares flash red (150ms each), however after the subject correctly selects Square1 and Square2 they must then select a blue triangle, which appears in one of the remaining locations. Subjects do not progress to the next training level until they achieve 80% accuracy, assessed after each 8 trial block.

After training, BASE and PROBE experimental trials were presented (Figure 1.6-1) and gaze data were recorded (sample rate = 30 Hz). For PROBE trials, the blue Triangle appears alongside the Square1 → Square2 demonstration and remains visible throughout subjects' response. Crucially, on PROBE trials, subjects can either continue to use the Square1→Square2→Triangle sequence [i.e., the learned strategy (**LS**)] *or* they can simply ignore the demonstration and select the Triangle [i.e., the direct strategy (**DS** or shortcut)]. BASE trials are identical to Training 3, however if subjects select the Triangle's 'hidden' location their response is recorded as if they had used the DS, providing a measure of accidental DS-use for each subject. In summary, the LS-DS task assesses subjects' propensity to forego their learned response and adopt the more efficient shortcut when it is available (PROBE trials), thereby measuring susceptibility to cognitive set.

Following the first 48 (PRE) experimental trials, consisting of 24 BASE and 24 PROBE trials randomly presented, subjects encountered a PAUSE screen. At which point, they were given a questionnaire requesting descriptions of various components of the task. Once they completed the questionnaire (~5-10 minutes), subjects were shown a brief video twice, demonstrating either the shortcut (Informed, n = 29) or the learned strategy (Control; n = 29) being performed in four consecutive PROBE trials. Groups were comprised of an equal number of males and females but were otherwise randomly assigned. After the video, subjects completed an additional 48 (POST) trials, again consisting of 24 BASE and 24 PROBE trials randomly presented, followed by another, identical questionnaire. Thus, each subject completed 48 PRE trials, a PRE questionnaire, 48 POST trials, and a POST questionnaire.[1]

4.2.4 Data Analysis

The percentage of correct trials in which the DS was used was calculated for both BASE and PROBE trials. For each subject, BASE DS-use was subtracted from PROBE DS-use to yield a measure of DS-use which accounts for within-subject error. In accordance with previous methodologies, we classified subjects that used the DS in greater than 5% of trials as DSers (Pope et al., 2015). Whenever the data violated the assumptions for parametric statistical analyses, non-parametric analyses were used and group medians rather than means were reported.

4.2.5 Gaze Analysis

For each trial, we assessed whether or not subjects looked at the Triangle prior to their first response. This was accomplished by classifying the raw data points from the eye tracker into Top Left, Top Right, Bottom Left, and Bottom Right quadrants, excluding the middle 50 pixels extending across the screen both vertically and horizontally, which also encompassed the central location of the fixation point. For each subject, the percentage of correct PRE trials in which they fixated on the Triangle's location before responding was calculated and compared between BASE and PROBE conditions. Remember, in BASE trials, the location of the Triangle is 'hidden' until Square 1 and Square 2 have been correctly selected; thus, this comparison controls for the number of times subjects might randomly look at the Triangle's location on the screen. Note, by only including data points in which subjects fixated, we excluded any potential saccades to the Triangle, making this a highly conservative measure.

4.2.6 Questionnaire Analysis

PRE and POST questionnaires were analyzed for indications that subjects recognized the Triangle's differential presence in PROBE and BASE trials. Thus, each subject received two scores of either noticed (1) or did not notice (0) based on their PRE and POST responses. Terms like "distraction" in reference to the Triangle were categorized as noticing a difference between BASE and PROBE trials and were scored as 1. Additionally, subjects' PRE and POST responses were analyzed for any mention of the Triangle as having either positive or neutral importance, compared with the squares. Subjects were scored based on whether they described the Triangle as more important (1) or of equal importance (0). Responses noting that it was "how to progress to the next trial," "how you knew you were correct," or "the goal" were considered indications of positive importance and were scored as 1. Thus, for both PRE and POST trial blocks, each subject received two scores, noticed/did not notice (1/0) and positive/neutral importance (1/0). The experimenter was blind to video condition during coding however, to ensure that bias did not factor into scoring, 24 (20.69% of the total) scores were re-coded by a second experimenter who was blind to trial block (PRE/POST) and the subjects' condition (Informed/Control). A Spearman rank order correlation between the two observers revealed that scoring was reliable ($r_s = .781$, $p < .001$).

4.3 Results

4.3.1 Effects of Seeing the Triangle

First, we investigated whether subjects saw the shortcut prior to their first response on PROBE trials. A Wilcoxon Signed-Rank test showed that indeed, subjects fixated on the Triangle's location more often prior to responding to PROBE trials (*M* = 22.93%, *SD* = 16.05) than BASE trials (*M* = 12.06%, *SD* = 8.43; *Z* = -4.34, $p < .001$; Figure 4.3-1). Figure 4.3-2 displays a heatmap of all subjects' gaze data for the Top Left → Bottom Left → Top Right configuration in both BASE and PROBE conditions.

Next, we explored whether visual search influenced strategy choice. If seeing the Triangle prior to the first response promotes DS-use, we would expect less LS-use following Triangle fixations. We chose to analyze LS-use because it is inversely related to DS-use, but is not conflated with whether or not subjects searched for the Triangle in order to select it (i.e., use the DS). We calculated the proportion of PROBE trials in which subjects used the LS after fixating on the Triangle (Look LS) or not (No Look LS). A Wilcoxon Signed Rank Test revealed that, for PRE trials, there was no difference between the proportions of LS-use depending on whether subjects did (*Mdn* = 100%; *SD* = 23.98) or did not (*Mdn* = 100%; *SD* =

17.48) see the Triangle (*Z* = -.051, *p* = .959), indicating that seeing the Triangle did not influence strategy selection for PRE trials.

4.3.2 Impact of Video Information

For POST trials, we analyzed the impact of the video information on visual search. A Mann-Whitney U test revealed that Informed subjects fixated on the Triangle significantly more often (*Mdn* = 45.83%, *SD* =5.15) than Controls (*Mdn* = 12.50%, *SD* = 9.23; U = 120.0 p < .001; see Figure 4.3-1) on PROBE trials. There was no significant difference between groups on BASE trials.

Further, after looking at the Triangle, Informed subjects (*Mdn* = 6.27%, *SD* = 45.40) used the LS less often than Controls (*Mdn* = 100.0%, SD = 36.25; *U* = 197, *p* < .001). Thus, after watching a demonstration of the DS, subjects were more likely to fixate on the shortcut *and* more likely to use it after seeing it.

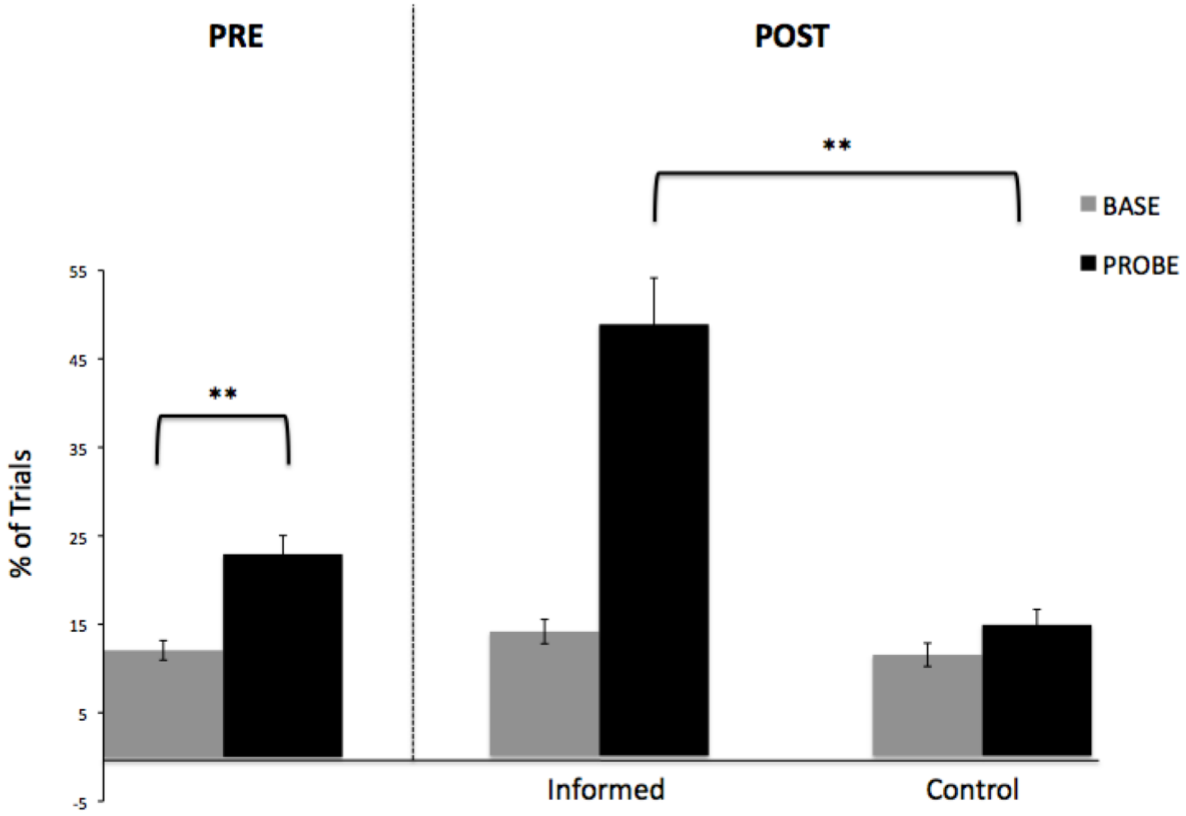


*Figure 4.1 The percentage of BASE and PROBE trials in which subjects fixated on the Triangle in PRE and POST blocks. Informed and Control subjects viewed the DS and LS video demonstrations, respectively. ** $p \leq .001$.*

Finally, to understand how the video information affected strategy-selection, we analyzed whether the proportion of DSers in PRE and POST trials differed between Informed and Control subjects. In other words, did the DS demonstration affect Informed subjects' strategy choices to a meaningful extent? For PRE trials, a Yates' continuity-corrected Chi square (two of the expected values were smaller than 5) confirmed that there was no difference in the proportion of Informed (13.8%) and Control (13.8%) DSers [$\chi^2(1, N = 58) = .000, p = 1.00$). For POST trials however, there were significantly more Informed (69.0%) than Control (6.9%) DSers [$\chi^2(1, N = 58) = 23.73, p < .001$; see Table 4.3-1].

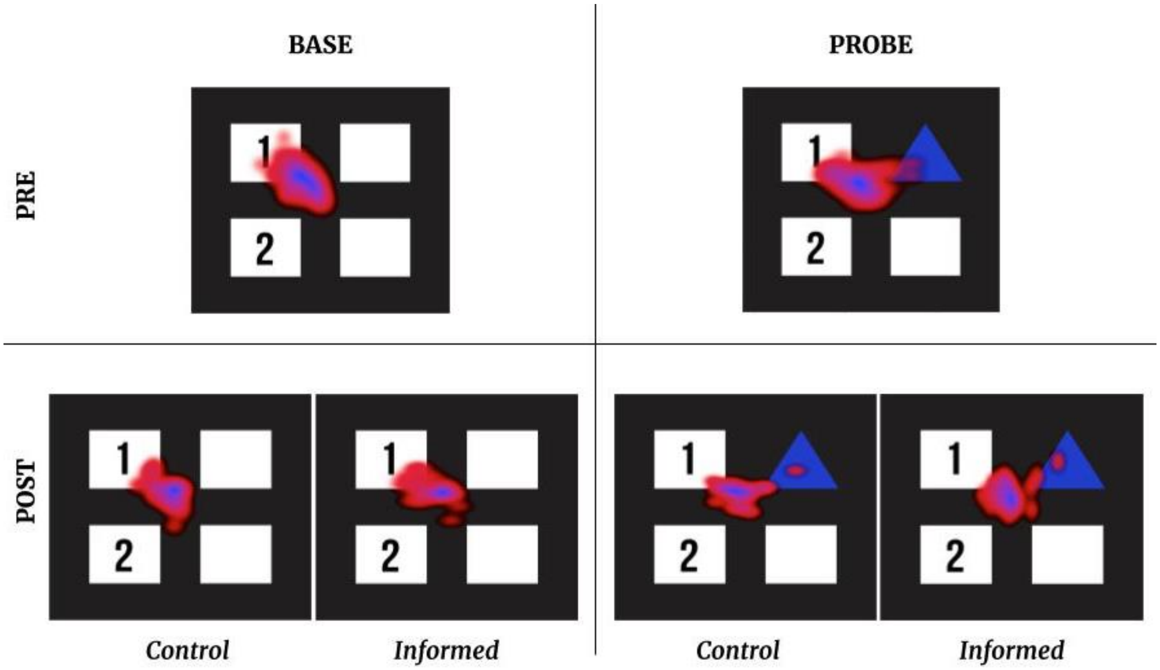


Figure 4.2 Response 1 gaze data compiled across all subjects during the Top Left (Square1), Bottom Left (Square2), Top Right (Triangle) BASE and PROBE trial configurations, for PRE and POST trial blocks.

We further explored the consistency of DS-use between groups by applying progressively more stringent DSer classification requirements: DS was used in more than 25%, more than 50%, and more than 75% of trials (Table 4.3-1). For PRE trials, as with the 5% classification criterion, there were no significant differences between groups. Additionally, only 2 subjects used the DS the first time it was available, both of whom went on to be categorized as DSers, and no subject used it in every PROBE trial. In contrast, for POST trials, Chi square analyses revealed that there were significantly fewer Control DSers than Informed DSers when the criterion was set to more than 25% [$\chi^2(1, N = 58) = 26.95, p < .001$], more than 50% [$\chi^2(1, N = 58) = 24.73, p < .001$], and more than 75% of trials, [$\chi^2(1, N = 58) = 20.62, p < .001$; Table 4.3-1]. Further, 14 subjects used the shortcut the first time it was available, 13 of whom had seen the

DS-video. All 14 went on to be classified as DSers and 4 of them used the DS every time it was available.

The task permits a third correct strategy, in which subjects begin to use the LS but then switch to the DS (Square1→Triangle). Subjects rarely used this ‘switch strategy’ in either PRE (*M* = 0.86, *SD* = 3.56) or POST (*M* = 1.81, *SD* =7.93) trials. However, a few subjects used it in more than 5% of PRE (*N* = 2, *Max* = 22.73%) and POST (*N* = 4, *Max* = 58.33%) PROBE trials.

4.3.3 Impact of Strategy on Performance

First, we isolated subjects’ first response time (i.e., the time between the end of the demonstration and their first response (RT1) for each trial. A Wilcoxon Signed-Rank test revealed that when the LS was used, subjects took longer to make their first response in PROBE (*Mdn* = 682.50 ms, *SD* = 190.09) compared to BASE (*Mdn* = 645.50, *SD* = 276.74) trials (*Z* = -2.139, *p* = .020).

We calculated switch costs for subjects who used the DS in greater than 50% of trials. We did not analyze PRE trials, as there were only four subjects who used the DS in greater than 50% of trials. However, for POST trials, a repeated measures ANOVA revealed that for greater than 50% DSers (*N* = 20) trials in which subjects repeated a strategy (e.g., LS-to-LS or DS-to-DS; stay trials; *M* = 614.48 ms, *SD* = 194.93) were significantly faster than trials in which subjects switched strategies [e.g., DS-to-LS or LS-to-DS; switch trials; *M* = 666.18, *SD* = 209.69; *F*(1,19) = 5.097, *p* = .036)]. There was no effect of BASE/PROBE condition. Additionally, a Wilcoxon Signed-Rank test confirmed that overall trial times were significantly longer when subjects used the LS (*Mdn* = 1605.90, *SD* = 430.85) compared to the DS (*Mdn* = 637.10, *SD* = 316.02; *Z* = -4.782, *p* < .001).

Table 4.1 Proportion of each group that used the DS in greater than 5%, 25%, 50%, and 75% of trials. Superscripts denote statistically distinguishable groupings.

	5%	**25%**	**50%**	**75%**
PRE				
Control	13.8^{a}	6.9^{a}	6.9^{a}	0^{a}
Informed	13.8^{a}	6.9^{a}	6.9^{a}	6.9^{a}
POST				
Control	6.9^{a}	3.4^{a}	3.4^{a}	3.4^{a}
Informed	69.0^{b}	69.0^{b}	65.5^{b}	58.6^{b}

4.3.4 Impact of Conceptual Understanding on Strategy

Next, potential interactions between subjects' questionnaire responses and their response strategies were assessed. For PRE trials, 46.6% of subjects (N = 27) indicated that they noticed a difference between PROBE and BASE trials and 24.1% of subject (N = 14) attributed positive importance to the Triangle. For PRE trials, Yate's continuity corrected Pearson's Chi-square tests showed that DSer classification was not significantly associated with either ascribing enhanced importance to the Triangle [$\chi 2(1, N = 58) = .26, p = .61$] nor noticing the difference between BASE and PROBE trials [$\chi 2(1, N = 58) = 1.84, p = .18$]. Further, 77.8% of subjects who noticed the difference between the PROBE and BASE trials and 50.0% of subjects who attributed increased importance to the Triangle still did not use the shortcut.

For POST trials, 37.9% of subjects (N = 22) reported noticing a difference between PROBE and BASE trials and 48.3% (N = 28) recognized the Triangle as being more important than the squares. A Yate's continuity corrected Pearson's Chi-square test showed that DSer classification was significantly associated with ascribing enhanced importance to the Triangle [$\chi 2(1, N = 58) = 13.88, p < .001$]. That said, 35.7% of subjects who attributed increased importance to the Triangle, still went on to use LS. Furthermore, DS-use was not significantly associated with noticing the difference between BASE and PROBE trials. The change between

PRE and POST appears to have been driven by Informed subjects, as only 2 Control subjects were classified as DSers in POST trials.

4.4 Discussion

In the current study we found that while operating under a cognitive set, subjects did, in fact, fixate on the shortcut prior to responding, yet seemingly ignored it. Consistent with this, in PRE trials, 27 out of the 58 subjects reported noticing a difference between PROBE and BASE trials and 14 subjects attributed greater importance to the Triangle over the squares; yet only 8 used the DS. After viewing a video demonstration of the DS however, more subjects fixated on the shortcut and then used it. In fact, prior to watching the video 13.8% of about-to-be Informed subjects used the DS, but this number rose to 69.0% after watching the video. Additionally, DS-use correlated with both noticing the difference between PROBE/BASE trials and ascribing increased importance to the Triangle but only in POST trials.

Our first hypothesis suggested that an inability to perceive the alternative strategy might have promoted cognitive set; however, this was not supported. Subjects clearly fixated on the shortcut yet subsequently used the LS. Additionally, when subjects used the LS, they took longer to respond to PROBE than BASE trials, further suggesting that the presence of the Triangle may have been distracting. Indeed, 77.8% of subjects who reported noticing the difference between BASE and PROBE trials still applied the LS to every trial indiscriminately. We conclude that cognitive set on the LS-DS task does not arise from an inability to visually perceive the shortcut.

Our findings are somewhat contradictory to those of Bilalić et al. (2008), who concluded that, in chess, subjects became stuck on a known solution, seemingly because they were not looking for alternatives. However, in chess, presumably the optimal move is no more visually salient than the familiar approach. The LS-DS task is designed to promote shortcut use by making the DS highly salient (during PROBE trials, the Triangle remains on the response screen amongst the white squares); thus, it is no surprise that subjects fixate on it. What the current study illustrates is that, even when subjects saw the alternative, they stuck to their familiar solution. Discovering the shortcut relied on subjects' propensity to explore the DS as an option after seeing it.

The current results support our second hypothesis that cognitive set is influenced by the ability to consider the alternative as a viable option. Informed subjects, who were explicitly shown that the DS could function as a possible solution, were 10x more likely to be classified as DSers than Control subjects, who saw the LS demonstrated. Thus, not surprisingly, understanding that they *could* use the shortcut greatly increased subjects' propensity to do so. However, even after watching the video, only 4 subjects used the DS every time it was available, illustrating subjects' inability to completely forsake their learned approach.

> *"When the individual does not adequately deal with problems but views them merely from the frame of reference of a habit; when he applies a certain habituated behavior to situations which have a better solution or which, in fact, are not even solvable by the just working habit; when, in a word, instead of the individual mastering the habit, the habit masters the individual – then mechanization is indeed a dangerous thing."* - *(Luchins, 1942, p. 93)*

Of note, a number of Informed subjects seemed wholly unaffected by the demonstration. Initially we piloted the use of extinction trials, wherein only the Triangle was available (i.e., no Square1→Square2 demonstration); however, these trials failed to invoke shortcut-use in subsequent trials. Indeed, subjects simply switched back to using the LS when the Square1→Square2 demonstration reappeared. Thus, we opted to directly demonstrate the DS, via the video. We expected that subjects would surely employ the shortcut after watching the video showing them how. Indeed, Informed subjects exhibited a stark increase in DS-use; however, 31.0% of subjects (*N* = 9/29) – even after watching the video demonstrating the shortcut, a minimum of 8 times – did not use it. Further, none of these nine subjects' POST questionnaire scores showed improvement over their PRE scores. Whether this arose out of an inability to extract information from, or simply inattention to, the video is unclear. Similarly, Luchins noted that if he explained the presence of the alternative strategy before the task began, subjects readily embraced it. However, even under these conditions, when the alternative was presented even a few trials after when subjects thought it would appear, a number of them did not use it (Luchins, 1942). This clearly highlights the trade-off between exploring other response options, which may take time and energy, and exploiting known solutions, which are effective but perhaps not the most efficient.

Like many tasks wherein subjects shift between multiple response strategies (Diamond & Kirkham, 2005; Ionescu, 2012; Lemaire et al., 2017; Luwel et al., 2009; Pope, Meguerditchian, Fagot, & Hopkins, In Prep; Stoet & Snyder, 2008), we noted an increase in the time it took subjects to respond when they switched between DS and LS responses. Chevalier, Blaye, Dufau, and Lucenet (2010) tracked subjects' gaze during a DCCS task and noted that switch costs were seemingly related to the time required to identify the newly relevant aspects of the stimulus.

Thus, it is possible that conservative LS-use might serve to mitigate the costs associated with switching strategies. Yet, the overall advantage of this approach is dubious, especially considering the increased efficiency of using the DS in overall trial times. Recall that the *Median* DS trial time was approximately 968.8ms faster than the *Median* LS trial time (See Section 4.3.3).

Interestingly, in PRE trials, 13.8% of Control subjects were classified as DSers; however, after watching the video demonstration of the LS, this number dropped to 6.9%. In other words, two DSers switched away from using the DS to embrace the LS as their primary PROBE response strategy, after viewing the LS video. This suggests that they were, in a sense, complying with the demonstration. Similarly, Luchins noted that, for his task, some subjects reported that they already knew about the shortcut but used the learned solution because they thought they were supposed to (Luchins, 1942, p. 89). Although we did not distinguish between subjects believing that they *could* not vs *should* not utilize the DS, future efforts should address the possibility that subjects' proclivity for rule-based responses may be driven by a desire to respond appropriately, rather than efficiently.

We conclude that cognitive set on the LS-DS task is not attributable to an inability to perceive the alternative and suggest that subjects' understanding of the problem space and willingness to try alternatives contributes greatly to their ability to break cognitive set. Additionally, we suggest that subjects' experience with rule-based problem solving might contribute to cognitive set. Alternative strategies are not used, because they are not sought. The impact of rote memorization and mechanized rule-use, typical of Western educational approaches, on cognitive inflexibility should be clearly elucidated in future endeavors.

Footnotes:

[1] We encountered several difficulties during data collection. First, occasionally the task would not resume following the PRE-questionnaire/video pause. When this occurred, the computer was restarted and a new program was launched which allowed subjects to return to the PAUSE screen (skipping Training and PRE trials); however, this process sometimes took several minutes. Therefore, subjects were shown the video once more before progressing to the POST trials. Second, for unknown reasons, the eye tracker stopped recording one subject's data on the last 10 POST trials. We chose to include this subject (*Informed*, age = 18, female) in our final analyses, which were all performed on proportions of trials (e.g., the proportion of correct trials, in which subjects used the DS). Finally, for another subject the computer froze just before the PAUSE screen. Thus, the experiment was restarted and the subject conducted an additional 48 trials (91 PRE trials total, rather than 48). Only the specific trials that were not completed in the first attempt were included in final PRE trial analyses.

5 ENHANCED COGNITIVE FLEXIBILITY IN THE SEMI-NOMADIC HIMBA

A version of this manuscript is being reviewed for publication in *Cognition*

Sarah M. Pope[1,2], Joël Fagot[2,3], Adrien Meguerditchian[2,3], David A. Washburn[4], William D. Hopkins[1,5]

[1] Neuroscience Institute and Language Research Center, Georgia State University, Atlanta, Georgia, USA

[2] Laboratory of Cognitive Psychology, CNRS and Aix-Marseille University, Marseille, France

[3] Brain and Language Research Institute, Marseille, France

[4] Language Research Center, Department of Psychology, Georgia State University, Atlanta, GA

[5] Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, Atlanta, GA, USA

Abstract

Through codified rule-use humans are able to accurately solve many problems; however, mechanized strategies can also be costly. After adopting a solution strategy, humans consistently become blind to alternatives – even when those alternatives are more efficient. This predisposition towards inflexibility, or cognitive set, was considered a universal phenomenon in humans however, prior to this study, only Western subjects had been tested. We used the nonverbal Learned Strategy-Direct Strategy (LS-DS) touch screen task in which subjects are presented with an opportunity to either use a learned strategy *or* a more efficient, but novel, shortcut. We found that the remote, seminomadic Himba of northern Namibia exhibited enhanced cognitive flexibility on the LS-DS task, thereby challenging the claim that cognitive set affects humans universally. Additionally, we did not find support for the notion that variation in local vs global perceptual processing contributes to cognitive flexibility. We discuss how other aspects of cultural variation, namely educational background, might contribute to the observed cross-cultural differences in problem solving flexibility.

5.1 Introduction

In 1942, Abraham Luchins asked a group of university students, faculty, and staff to solve a set of simple math problems (Luchins, 1942). He aimed to explore how rule-based problem solving limits other, more creative solutions. The task began with several problems solvable only by using a four-step rule, which participants quickly mastered. However, after these 'learned rule' problems, Luchins added a twist. In addition to the familiar strategy, problems seven and eight could also be solved by a more efficient, one-step method – a shortcut. Remarkably, when the time came, not a single subject used the shortcut. Instead, they continued to use the learned rule despite its relative inefficiency. This phenomenon - the inability to implement an alternative strategy once a learned rule has been adopted - is termed 'cognitive set.'

Pope et al. (2015) devised a nonverbal, nonmathematical adaptation of Luchins' task in order to compare baboons' and humans' susceptibilities to cognitive set (Pope et al., 2015). The LS-DS task begins with several levels of training wherein subjects learn to utilize a three-step (Square1→Square2→Triangle) sequence, which constitutes the learned strategy (**LS**; see Figure 1.6-1b). Once subjects consistently utilize the LS, experimental trials are presented in which subjects can use the LS *or* they can use a direct strategy (**DS** or the shortcut) by skipping Square1 + Square2 and going straight for the Triangle (See Figure 1.6-1d). We found that all of the 15 baboon subjects immediately switched to the DS when it became available and used it in 99.9% of trials. Similar to Luchins' findings, only 4 out of the 53 (i.e., 7.6%) adult human subjects used the DS in more than 5% of trials. That is to say, adult humans but not baboons were affected by cognitive set on the LS-DS task.

Flexible problem solving is a critical element of navigating dynamic environments; hence, its role in cognition has been extensively studied. Typical cognitive flexibility measures, such as Discrimination Reversal, Transfer Index, and Wisconsin Card Sorting tasks, first allow a subject to learn a correct solution method and then require them to abandon and replace it with another (Berg, 1948; Harlow, 1949; Rumbaugh, 1971). Flexibility is then calculated as the degree of persistence before the new, correct solution strategy is adopted. In these and other cued task-switching paradigms, in which subjects switch strategies repeatedly, pronounced 'switch costs' have been noted in both children and adults (Diamond & Kirkham, 2005; Ionescu, 2012; Lemaire et al., 2017; Luwel et al., 2009; Stoet & Snyder, 2008). Switch costs are deficits in response time or accuracy following a strategy switch and are thought to be associated with disengaging with one strategy and initiating another (Meiran, 1996; Rogers & Monsell, 1995; Stoet & Snyder, 2007); however they are not exhibited by some nonhuman primates (Stoet & Snyder, 2003; however, see Huguet et al. 2014), suggesting that they may be a byproduct of human rule-encoding (Stoet & Snyder, 2008).

Cognitive set tasks, like Luchins' and the LS-DS task, differ from typical measures of cognitive flexibility in that subjects are not *required* to switch strategies. Subjects can (and do) continue to use the learned strategy and it continues to be rewarded, albeit at a less efficient rate than if they are flexible and switch to the alternative. We argue that this may be more representative of flexibility in real life, where problems commonly have multiple solutions from which to choose. Since it's debut, thousands of subjects have been tested on Luchins' task with various manipulations and consistently, 70-100% follow the learned rule (Luchins, 1942; Luchins & Luchins, 1950). Luchins (1942; 1950) and others (McKelvie, 1984) concluded that within human problem solving, this propensity toward cognitive set is universal. However, although attempts were made to account for subjects' age, occupation, and education, any potential cross-cultural differences in susceptibility to cognitive set were completely neglected.

Recently, numerous studies have called into question the 'universality' of cognitive processes, our understanding of which has been predominantly derived from Western, Educated, Industrialized, Rich, Democratic (WEIRD) subjects' data (Henrich et al., 2010). These findings illustrate cultural differences in cognitive processing based on both visual perception (Caparos, Ahmed, et al., 2012; Caparos, Linnell, et al., 2012; Davidoff et al., 2008; de Fockert et al., 2011; de Fockert et al., 2007; Fagot, Goldstein, Davidoff, & Pickering, 2006; Ji et al., 2000; Masuda & Nisbett, 2001; Nisbett et al., 2001; Nisbett & Miyamoto, 2005) and problem conceptualization (Abel & Hsu, 1949; Berl & Hewlett, 2015; Chiu, 1972; Clegg & Legare, 2016; Frank et al., 2008; Ji et al., 2000; Legare & Nielsen, 2015; Nisbett et al., 2001). Moreover, in typical (and arguably less indicative) measures of cognitive flexibility such as the WCST, some cross-cultural differences have been found, yet have been attributed to flaws in the task itself rather than potential variation in cognitive processes across cultures (Avila, 2013; Coffey et al., 2005; Shan et al., 2008). In fact, our own previous findings noted an effect of age on cognitive flexibility; specifically, children 7-10 were more likely to use the shortcut than adolescents (11-18) or adults (19-68) in the LS-DS task (Pope et al., 2015; however, see Chapter 3 Footnote 2), a further indication that cognitive set does not affect humans uniformly.

As we have shown before, baboons, which process visual information with a local bias (Deruelle & Fagot, 1998) are better able to break cognitive set and use the shortcut than Western adults (Pope et al., 2015), who process visual information with a global bias (Navon, 1977). This led us to consider that differences in global versus local visual processing between humans and baboons might explain their divergent performance in breaking cognitive set. Namely, perhaps humans' attention to the global similarities between the LS-DS trials might have promoted the use of the learned rule (i.e., cognitive set) even when relevant, local details (i.e., the presence of the Triangle) were present that could have lead them to the shortcut (which the baboons found and used immediately). Note, this is in stark contrast to the literature purporting that global/local processing biases stem from holistic/analytic processing styles, that result in flexible/persistent problem-solving strategies, respectively (Hommel & Colzato, 2017; Ji et al., 2000; Masuda & Nisbett, 2001, 2006; Nisbett et al., 2001). Granted, there is limited support for a causal link between a global mindset and global perceptual processing (Kuhnen & Oyserman, 2002), however the assertions regarding perceptual processing and problem-solving approach are largely unsupported (Davidoff et al., 2008; Peterson & Deary, 2006). Instead, we suggested that a visual inability to discriminate the DS from the sequence as a whole might factor greatly into humans inability to utilize the shortcut. An example of this inadvertent blindness comes from previous research showing that after finding a non-ideal solution, chess players reported that they were looking for alternatives, yet their eye movements did not stray from the one they had already found (Bilalić et al., 2008).

Alternatively, differences in problem conceptualization might have driven the divergent susceptibilities to cognitive set between baboons and humans. Namely, human subjects may have responded according to how they thought they *should*, based on their previous experience with rule-based problem solving. Undoubtedly, rules-of-thumb and equations offer a tremendous advantage when solving many, similar problems. For this reason, repetitive rule-use is a staple of Western education, yet the real-world applications of this problem solving approach might reasonably lead to cognitive inflexibility (Star & Seifert, 2006). Alternative solutions are not found because they are not sought. The LS-DS task does not instruct subjects how to respond, they learn through trial and error however, it is possible that previous human subjects

(Pope et al., 2015) did not consider the DS a viable option based on their history with repetitive rule-use and/or single-solution problems.

The current study aimed to 1) test the hypothesis that cognitive set is a universal human trait, invariant to cultural factors and 2) determine the relative effects of perceptual and conceptual influences on susceptibility to cognitive set. The seminomadic Himba of northern Namibia exhibit a strong local perceptual bias compared to Westerners (Davidoff et al., 2008; de Fockert et al., 2007). Thus, we investigated the differences in cognitive set between Westerners and traditional Himba using the LS-DS task (Pope et al., 2015). If a global perceptual bias governed the previously tested Western humans' susceptibility to cognitive set, then traditional Himba should be better able to use the shortcut than Westerners. Yet, how could we confidently attribute this to their local perceptual processing, rather than broader cultural distinctions between traditional Himba and Westerners? Opportunely, a subpopulation of Himba have moved to the small nearby town of Opuwo. Previous research has shown that traditional Himba are less susceptible to the Ebbinghaus illusion and exhibit greater selective attention/reduced distractibility than both Urban Himba and British participants, who responded similarly to one another (Caparos, Ahmed, et al., 2012; Caparos, Linnell, et al., 2012; de Fockert et al., 2011; Linnell et al., 2013). These differences in perceptual processing between Urban and Traditional Himba provide a natural control for external factors such as genetics, religion, language, climate, and education. Thus, if perceptual processing affects shortcut-use, we predicted that urban Himba would respond to the LS-DS task in a similar manner as Westerners but different from traditional Himba.

In consideration of the conceptual influences hypothesis, we reasoned that if subjects' conceptualization of the task affects cognitive set, then releasing subjects from potentially constraining presuppositions would enhance shortcut-use. In Luchins' task, he achieved a degree of success by using the prompt "Don't be blind" (Luchins, 1942). However, he also noted that some subjects interpreted this as *Don't be blind to the obvious rule*. We decided to give the prompt: "Don't be afraid to try new things" halfway through the LS-DS task, predicting that subjects' shortcut use would increase following the prompt. Furthermore, any relative differences in the prompt's influence on shortcut-use between Westerners, Traditional Himba, and Urban Himba might be indicative of the extent to which they were initially constrained by preexisting conceptualizations.

5.2 Methods

5.2.1 Subjects

Data were collected from 54 Western (42 females and 12 males), 54 urban Himba (34 females and 20 males), and 75 traditional Himba subjects (34 females and 41 males). All subjects were above 18 years of age; however, exact ages were not recorded because Himba do not keep an accurate record of their age. All methods were approved by the Georgia State University Institutional Review Board prior to testing. Western subjects were recruited from the pool of undergraduate students at Georgia State University by posting the study on the SONA Experiment Managements System, tested on the Georgia State University campus, and received one course credit in exchange for their participation.

Himba subjects were recruited and tested in their villages by a hired guide after initial permissions were received from the village leader(s). Testing took place inside a tent or in a shaded area. Himba subjects were classified as 'Urban' if they lived within 20km of Opuwo, the primary city in the Kunene region and 'Traditional' if they lived further than 100km from

Opuwo. We only tested subjects that fell into these two categories. The number of times each subject had been to a city (almost exclusively Opuwo) was recorded for all Himba subjects. Urban subjects received 20 Namibian Dollars and Traditional subjects received 1kg maize meal and 1kg sugar each.

5.2.2 LS-DS Task

The LS-DS task was programmed with OpenSesame software (Mathot et al., 2012) and administered via a Lenovo Ideapad FLEX 4 (14”) 2-in-1 touchscreen laptop. All subjects received basic instructions on touching the fixation cross to begin each trial and which feedback screens/sounds indicated correct vs incorrect responses. As part of the instructions, all subjects were told that they would need to touch the shapes to figure out the correct answer. However, in the Himba’s language (Otjihimba) there is not a direct translation for ‘shapes.’ Thus, Himba subjects were shown an illustration of a square and a triangle at that point during the instructions.[1]

A complete description of the LS-DS task can be found in (Pope et al., 2015); however, briefly: the task consists of three training levels and 96 experimental trials. In Training 1, a demonstration shows two squares, which flash red sequentially (250ms each); the subject must reproduce the demonstration by selecting the two squares in the correct order. In Training 2, the demonstration shows four squares total, two of which flash red sequentially (200ms each) and the subject reproduces the demonstration by selecting the two correct squares (now out of four options) in the correct order. Training 3 is identical to Training 2 however after the subject has correctly selected Square1 and Square2 (which are demonstrated for 150 ms each), they must then select a blue triangle which appears in one of the remaining locations. Subjects do not progress to the next training level until they achieve 80% accuracy, measured after each 8 trial block. After training, BASE and PROBE experimental trials are presented (Figure 1.6-1a-d). In PROBE trials, the Triangle appears alongside the Square1 → Square2 demonstration and remains visible on the response screen. Importantly, to be correct subjects can either continue to use the Square1→Square2→Triangle sequence (i.e., the **LS**) *or* they can simply ignore the demonstration and select the Triangle (i.e., the **DS** or shortcut). BASE trials look identical to Training 3 however if subjects select the Triangle's hidden location they are marked as having used the DS; thereby providing a measure of accidental DS-use within each subject. In summary, subjects initially learn via the training to reproduce a two-square demonstration and then touch the Triangle. Cognitive set is measured by their propensity to forego this learned method in order to use a more direct strategy when the Triangle is already present.

For the current study, after the first 48 PRE trials, the task was paused and the experimenter issued the "Don't be afraid to try new things" prompt. Subjects then completed an additional 48 POST trials. PRE and POST trial lists consisted of one BASE and one PROBE trial for each of the possible 24 configurations; trial order was randomized.

5.2.3 Data Analysis

The percentage of accurate trials in which subjects used the DS was calculated for both BASE and PROBE trials. Each subject's BASE DS-use was subtracted from PROBE DS-use to yield a measure of DS-use that accounts for within-subject error. In accordance with our previous studies, subjects were classified as DSers if they used the DS in more than 5% of trials however, we also included progressively more stringent DSer qualification criteria (more than 25%, more than 50%, and more than 75%) to gain a more complete understanding of DS-use between groups. Whenever the data violated the assumptions for parametric statistical analyses, we used non-parametric analyses and reported group medians rather than means.

The switch costs associated with using the DS were also analyzed. Recall that during BASE trials, subjects typically only use the LS. Thus, for DSers who were switching between the DS and the LS for PROBE and BASE trials respectively, we expected to see switch costs. We isolated the time between the end of the Square1→Square2 demonstration and subjects' first response (RT1) for BASE trials in which subjects repeated (BASE *stay*) or switched (BASE *switch*) their strategy and for PROBE trials in which subjects repeated (PROBE *stay*) or switched (PROBE *switch*) their strategy. All trials that were precluded by an incorrect trial and those in which the first response was incorrect were excluded. Finally, we applied a natural log transform to normalize the data. Because some subjects did not ever use the DS, and thus did not switch, we analyzed only those who had used the DS in more than 50% of PRE (n=40) and POST (n=50) trials. Group (Western, Urban, Traditional) did not significantly impact any of the analyses.

5.3 Results

5.3.1 DS-Use

In accordance with our previous studies, subjects were classified as DSers if they used the DS in more than 5% of trials. A Pearson's chi-square revealed that, in PRE trials, a significantly smaller proportion of Westerners were classified as DSers (5.6%; n = 3) than either Urban (42.6%; n =23) or Traditional (34.7%; n = 26) Himba participants [$\chi^2(2, N = 183) = 20.65$, $p < .001$], whose proportions did not differ from each other (Figure 2).

5.3.2 Conceptual Influences on Cognitive Set

In POST trials, after subjects were told: “Don’t be afraid to try new things,” a Pearson’s chi-square analysis revealed that the proportion of DSers significantly increased for Westerners (50.0%; n = 27), significantly decreased for Urban Himba (27.8%; n = 15) and did not change for Traditional Himba [36.0%; n = 27; $\chi^2(2, N = 183) = 38.989$, $p < .001$]. In other words, the prompt effectively increased DS-use for Western, but not Himba, participants (Figure 5.3-1).

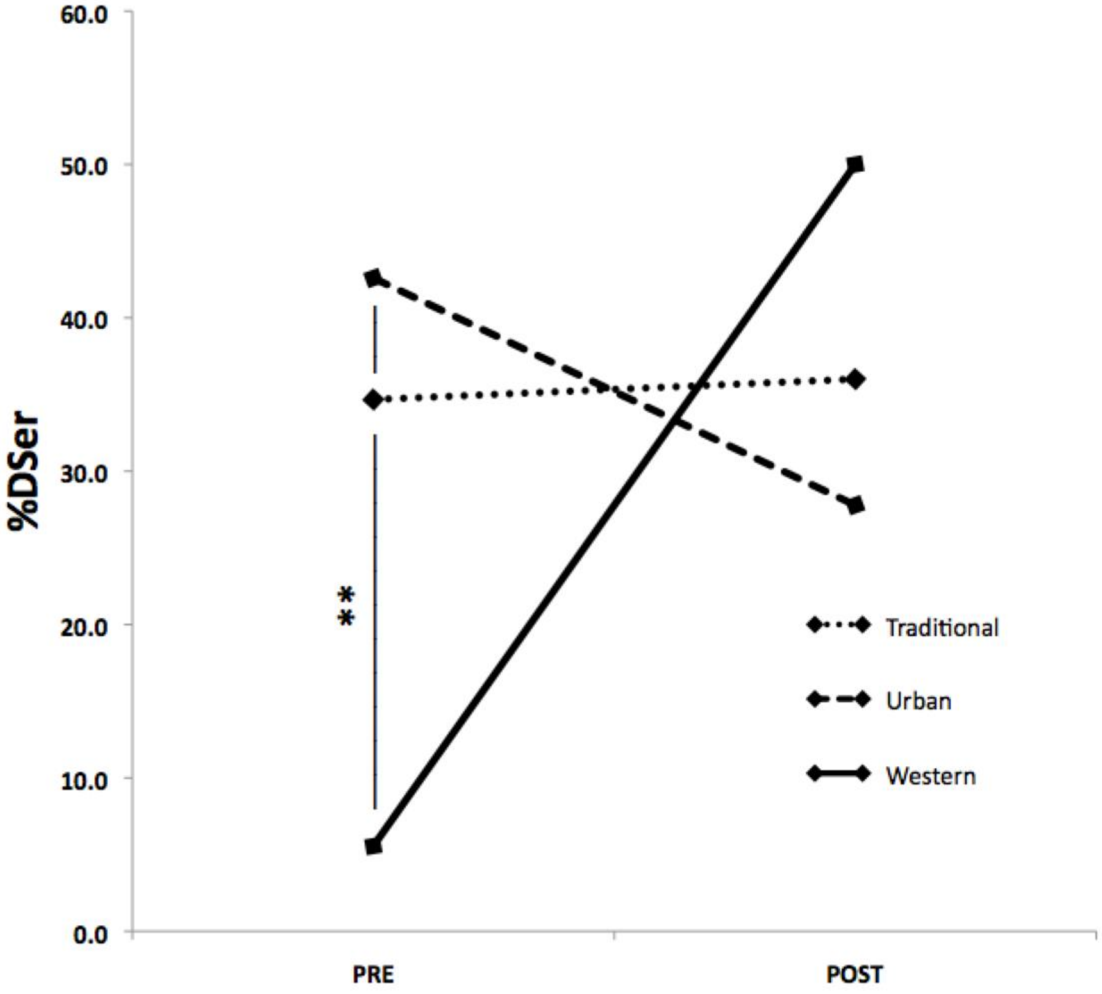


*Figure 5.1 The proportion of DSers within each group for before (PRE) and after (POST) receiving the prompt. *** p ≤ .001.

5.3.3 *Immediacy and Consistency of DS-use*

Significantly more Traditional (25.3%) and Urban (27.8%) Himba used the DS the very first time it was available (i.e.,, in PRE trials) compared with Westerners [3.7%; $\chi^2(2, N = 183) = 12.48, p = .002$]. However, following the prompt (i.e., in POST trials), Westerners (24.1%) used the first available shortcut as much as Urban (20.4%) or Traditional (33.3%) Himba [$\chi^2(2, N = 183) = 2.976, p = .226$]. We next investigated whether any subjects used the DS in every single trial. In PRE trials, 3.7% of Westerners, 3.7% of Urban, and 9.7% of Traditional Himba used the DS every single time it was available and, in POST trials, those numbers increased slightly to 5.6% of Westerners, 5.6% of Urban, and 18.7% of Traditional Himba. Notably, of the three Westerners classified as DSers in PRE trials, two of them used the DS in every trial that it was available and the third used it in 95.3% of trials, illustrating a stark contrast between them and Western LSers.

We decided to expand our understanding of the consistency of DS-use across groups by applying progressively more stringent DSer classification requirements: DS was used in more than 25%, more than 50%, and more than 75% of trials (Table 5.3-1). For PRE trials, as with the 5% classification criterion, there were significantly fewer Western DSers than Traditional or Urban Himba DSers when the criterion was set to more than 25% of trials [$\chi^2(2, N = 183) = 14.57, p = .001$], and more than 50% of trials [$\chi^2(2, N = 183) = 13.07, p = .001$]. When the criterion was set to more than 75% of trials, there were significantly more Traditional Himba DSers than Western DSers, but Urban Himba did not significantly differ from either group [$\chi^2(2, N = 183) = 9.099, p = .011$]. In POST trials, DSer proportions did not significantly differ between groups for any of the DSer classification criteria.

Table 5.1 Proportion of each group that used the DS in greater than 5%, 25%, 50%, and 75% of trials. Sub-groups based on the number of city visits are included for Traditional Himba. Superscripts denote statistically distinguishable groupings.

PRE	**>5%**	**>25%**	**>50%**	**>75%**
Traditional	34.7^{a}	33.3^{a}	32.0^{a}	25.3^{a}
0-1 visits	39.0^{a}	36.6^{a}	34.1^{a}	26.8^{a}
2+ visits	29.4^{a}	29.4^{a}	29.4^{a}	23.5^{a,b}
Urban	42.6^{a}	29.6^{a}	24.1^{a}	14.8^{a,b}
Western	5.6^{b}	5.6^{b}	5.6^{b}	5.6^{b}
POST				
Traditional	36.0^{a}	32.0^{a}	30.7^{a}	26.7^{a}
0-1 visits	41.5^{a}	36.6^{a}	36.6^{a}	29.3^{a}
2+ visits	29.4^{a}	26.5^{a}	23.5^{a}	23.5^{a}
Urban	27.8^{a}	22.2^{a}	20.4^{a}	20.4^{a}
Western	50.0^{a}	42.6^{a}	29.6^{a}	20.4^{a}

5.3.4 *Influence of Urban Exposure*

In Caparos et al.'s assessment of the differences between Westerners and Urban and Traditional Himba's perceptual processing, they noted that Traditional Himba that had visited the city fewer times within their lifetimes categorized more figures based on local, rather than global, features (Caparos, Ahmed, et al., 2012); thereby suggesting that that exposure to an urban environment corresponds to lasting changes in perceptual processing, biasing global rather than local features. To explore how urban exposure might have impacted our assessment of LS-DS performance, we re-grouped the Traditional Himba into those that had visited the city either one or fewer (*N* = 41; *M* = 0.24; *SD* = 0.43) or two or more times (*N* = 34; 16 participants reported "many" instead of an exact number, but for the other 18 Traditional Himba subjects, *M* = 3.06; *SD* = 1.70). In PRE trials, for the more than 5% DSer classification criterion, a Pearson's chi square analysis again revealed that the proportion of DSers within the Western group was significantly less than the Himba groups' [$\chi^2(3, N = 183) = 21.49$, $p < .001$; Western = 5.6%, Urban = 42.6%, Traditional 0-1 = 39.0%, and Traditional 2+ = 29.4%]; however, there was no difference in the proportion of DSers between Himba based on location (Urban vs Traditional) or the number of urban exposures (Table 5.3-1). This finding was preserved for the more than 25% and more than 50% DSer classification criteria. However, for the more than 75% DSer classification criterion, only Traditional Himba that had been to the city 0 or 1 times were statistically distinct [$\chi^2(3, N = 183) = 9.25, p = .026$] from Westerners, and both Traditional Himba that had been to the city 2+ times and Urban Himba displayed intermediate DS-use (i.e.,, the proportions of DSers within these groups did not differ significantly from either of the other groups or each other). Like before, in POST trials, DSer proportions did not differ significantly between groups for any of the DSer classification criteria.

5.3.5 Effect of Training

As noted previously, Western participants were more likely to successfully complete the training than Himba, who often took much longer to reach criterion. A Kruskal-Wallis analysis of variance confirmed that Westerners differed from Urban and Traditional Himba (which did not differ from each other) in the total number of training trials needed before graduating to the experiment [$H(2) = 63.48, p < .001$; *Mean* trials (*SD*): Western = 39.0 (33.44), Urban = 73.2 (40.12), Traditional = 86.2 (44.98)]. To assess whether increased training affected strategy-use, we used Logistic Regressions to analyze the effects of group (Western, Urban, Traditional) and # of training trials on DSer classification. As noted before, group significantly corresponded to DSer classification, correctly predicting 71.6% of cases. However, there were no significant effects of training (for any level) on DSer classification.[2]

5.3.6 Effects of Shortcut-Use on Performance

To explore the impact of strategy on accuracy, subjects' PROBE trial accuracies were compared using a Mann-Whitney U test, which indicated that for PRE trials, DSers (*Mdn* = 91.7%) had significantly higher (*U* = 4,187, *p* = .015) accuracy scores than LSers (*Mdn* = 83.3%). This was also true for POST trials, DSers (*Mdn* = 95.8%) had higher accuracy than LSers (*Mdn* = 83.3%; U = 6,202, *p* <.001). Next, to investigate the impact of strategy on efficiency, we compared subjects' average trial times for DS and LS responses in PROBE trials. A natural logarithm transform was applied to normalize the data. A paired samples T-test revealed that trial times were significantly faster when subjects used the DS (*M* = 1655.1 ms; *SD* = 1194.0) than when they used the LS (*M* = 2543.6; *SD* = 1459.0), validating the enhanced efficiency of the DS; t(84)=8.13, *p* <.001.

5.3.7 Switch Costs

We also identified the switch costs for subjects who utilized both DS and LS solutions. For PRE DSers, a repeated measures ANOVA revealed a significant (*F*(1,39) = 5.566, *p* =.023) interaction between condition (BASE vs PROBE) and trial type (*stay* vs *switch*) on subjects' RT1s. PROBE *stay* RT1s (*M* = 1188.6 ms; *SD* = 540.2) were significantly faster than PROBE *switch* (*M* = 1373.4 ms; *SD* = 816.3), which were faster than both BASE *stay* (*M* = 1686.0 ms; *SD* = 961.1) and BASE *switch* (*M* = 1587.5 ms; *SD* =637.1) RT1s, which did not differ from each other. Thus, although DS responses were faster than LS responses, DSers experienced switch costs during their LS-to-DS strategy switch.

For POST trials, again, we found a significant [*F*(1,49) = 4.56, *p* = .038] interaction between condition (BASE vs PROBE) and trial type (*stay* vs *switch*). Again, PROBE *stay* RT1s

(*M* = 850.2 ms; *SD* = 472.3) were significantly faster than PROBE *switch* (*M* = 960.7 ms; *SD* = 495.3) RT1s, which were faster than both BASE *stay* (*M* = 1233.9 ms; *SD* = 543.6) and BASE *switch* (*M* = 1261.2 ms; *SD* =532.8) RT1s, which did not differ from each other. Thus, the DS was, in fact, the more efficient strategy, even in the face of switch costs associated with going between LS and DS responses for BASE and PROBE trials, respectively.

Lastly, we assessed LSers' RT1s for BASE and PROBE trials. For PRE trials, BASE RT1s (*M* = 1152.2 ms, *SD* = 535.1) were significantly faster than PROBE RT1s (*M* = 1297.7 ms, *SD* = 600.7; t(142)=-10.95, *p* <.001). This was also found for POST trials (BASE *M* = 1148.9 ms, *SD* = 440.1; PROBE *M* = 1202.6 ms, *SD* = 542.5; *t*(132)=-2.19, *p* <.031) suggesting that the presence of the Triangle in PROBE trials may have been distracting.

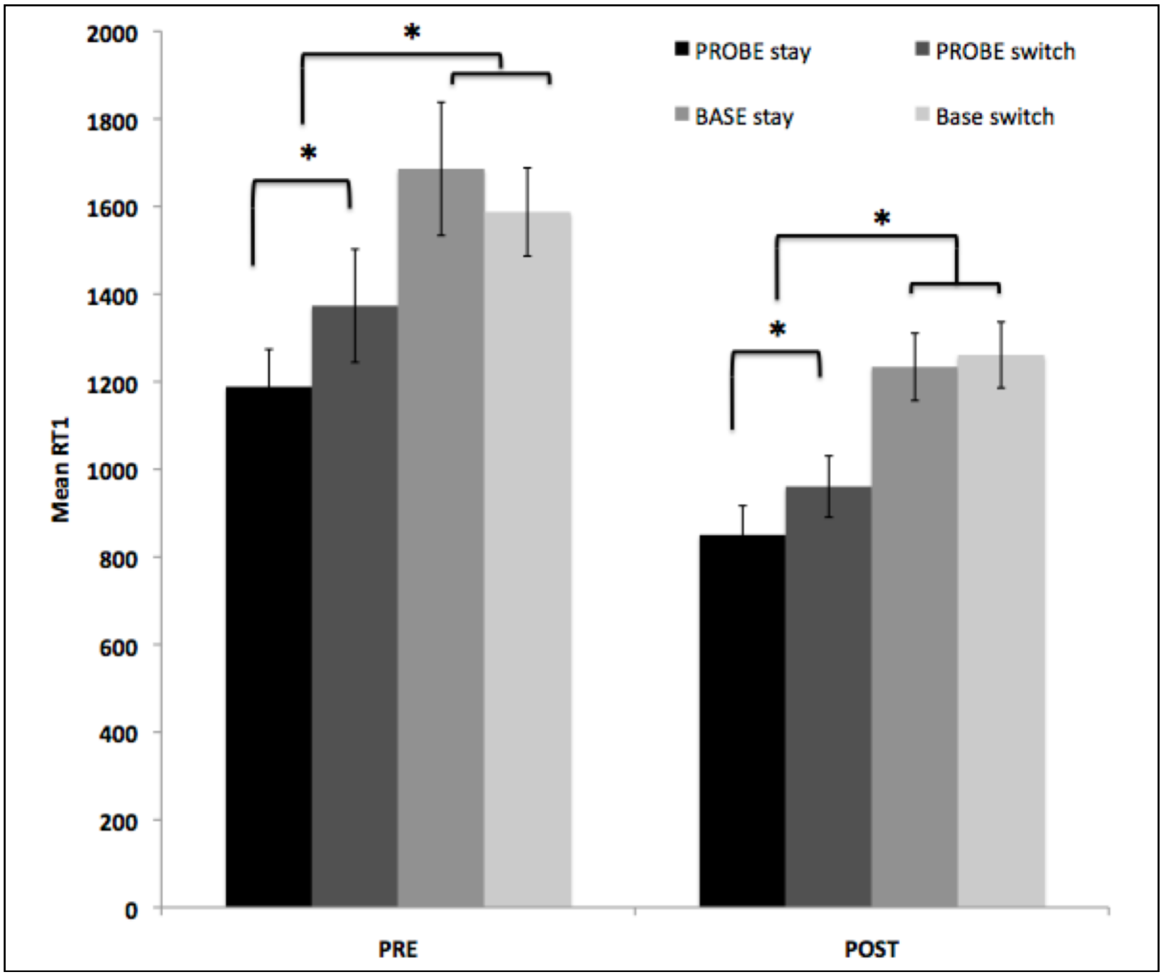


*Figure 5.2 Mean BASE and PROBE same and switch trials' RT1s for subjects who used the DS in greater than 50% of trials. * < .05.*

5.4 Discussion

The ability to break away from a learned rule and use a shortcut was enhanced in Urban and Traditional Himba compared to Westerners, providing the first evidence that cognitive set is not a universal human trait. Our findings do not support the hypothesis that differences in global vs local perceptual processing affect susceptibility to cognitive set on the LS-DS task, based on the previously observed differences between Urban and Traditional Himba's perceptual processing (Caparos, Ahmed, et al., 2012) and their predominantly indistinguishable strategy selection on the LS-DS task. Our efforts to promote shortcut-use, by issuing the "Don't be afraid to try new things" prompt, successfully increased the proportion of Western DSers to levels equivalent with the Himba. These findings support the hypothesis that cognitive set is influenced by participants' conceptualization of the problem. That being said, a large proportion of Westerners (50%; n = 27), Urban (72.2%; n = 39), and Traditional (64%; n = 48) Himba participants were still affected by cognitive set in POST trials: they did not adopt the shortcut even after the prompt. We discuss this persistence of cognitive set and propose potential causes, repercussions, and solutions.

The Himba's enhanced DS-use provides a clear contrast to Westerner's firm LS-use, allowing us to glimpse potentially causative environmental factors which might be limiting problem solving flexibility in Westerners. These two culturally distant groups differ in many ways, including social structure, physical environment, language, and educational background. Although, some of these factors have qualities more likely to influence cognitive flexibility than others.

First, a consequential impact of social structure on shortcut-use seems unlikely when one considers that male and female Himba play vastly different social roles, yet we found no sex-differences in LS vs DS strategy selection (see supplementary data in Appendix B). Further, the social structure hypothesis predicts that interdependent (such as the Himba) more than independent peoples (such as Westerners) should utilize more holistic response styles (i.e., consider the sequence as a whole (Nisbett et al., 2001); however, we found the opposite. Compared to Westerners, the Himba were better able to extract the Triangle from the larger LS sequence.

Second, the possibility that physical environment or visual clutter might play a role in susceptibility to cognitive flexibility also seems unlikely, given that we found very little evidence that urban exposure biased individuals' response strategy, as would be predicted by the visual clutter hypothesis. Traditional Himba that had been to the city two or more times and Urban Himba did exhibit intermediate DS-use but only when the criterion was set to more than 75% of trials. This is indicative of a potential association between urban exposure and the *consistency* with which subjects evoked the shortcut, rather than their propensity to break cognitive set in the first place.[3]

Third, language, or verbal encoding, is considered vital to rule-based problem solving (Stoet & Snyder, 2008). Without a word for 'shapes' the Himba may have been less equipped to verbally encode the LS, resulting in it imposing relatively weaker constraints on problem solving compared to Westerners. This could also explain the comparatively lower number of Himba participants who passed the training levels. We suggest that weaker verbal encoding would be a more viable explanation if *all* of the Himba subjects were less susceptible to cognitive set, given that *all* of the Himba were monolingual in Otjiherero. Yet, there were fewer DSers than LSers, even in Himba subjects. Further, our finding that Himba and Westerners were equally affected by switch costs suggests that the rule is similarly encoded in both populations. That said, the possibility that the strength with which the initial habit is codified could influence shortcut-use should be explored in future studies.

Consequently, we are left to focus on the putative impact of educational background (i.e., problem solving history) on cognitive set. Luchins' discussed the potentially set-inducing instructional methods typical of Western education in his initial description of cognitive set (Luchins, 1942). He asserted that, "Methods are needed which will teach the child to stand on his own feet, to face the world freely and act through intelligent thinking rather than by blind force of habit" (p. 93). We posit that the blind repetition characteristic of Western education deters subjects from interpreting the DS as a viable solution. They may never even consider that the problem could have multiple solutions, until explicitly told, "Don't be afraid to try new things" which clearly states the possibility of multiple solutions.

In support of this, following the prompt, 24.1% of Westerners (compared to 3.7% before) used the shortcut the very first time it was available. Additionally, we noted that Urban Himba's shortcut use significantly *decreased* following the prompt, however considering that in PRE trials Urban Himba showed the largest degree of shortcut-use, this might be interpreted as adherence to the prompt's suggestion to try new things. We argue that, because the prompt did not elicit enhanced shortcut use in either Urban or Traditional Himba, they may have already been operating without a single-solution assumption. Instruction has been found to induce set in other paradigms (Chrysikou & Weisberg, 2005) and this susceptibility may even vary across cultures (Berl & Hewlett, 2015; Clegg & Legare, 2016; Legare & Nielsen, 2015). A small caveat to this interpretation is necessary: although the vast majority of Himba do not partake in formal education, a very small portion of both Urban and Traditional Himba attend or have briefly attended small schools. We estimate that very few (<10%) of our subjects had been exposed to schooling and posit that any potential influence would have served to suppress DS-use, not enhance it.[4] Nevertheless, future studies directly aimed at addressing the impact of educational background and literacy on cognitive flexibility, especially in remote cultures, would be sapient.

In this study we confidently reject the universality of cognitive set by demonstrating a remote culture's enhanced ability to break away from a set strategy to adopt a more efficient alternative. We effectively promoted shorctut-use in a population previously bound to a learned rule by altering their conceptual understanding of the task. Lastly, we proposed that educational background might contribute to the observed cross-cultural differences in cognitive set. Simply put, we may be mechanizing ourselves by approaching multivalent problems with single-strategy solutions.

As with any cognitive process, cognitive set is likely mediated by multiple influences. This should certainly be taken into consideration when studying cognitive processes, not only to avoid erroneous predictions but also in an effort to understand why certain differences arise. What exactly allows some participants but not others to readily switch to the shortcut remains opaque but could be vital to enabling the rest of us to intentionally utilize both persistence and flexibility in problem solving. As a final consideration, we highlight the large proportion of LSers even during each group's peak DS-use (Westerners POST: 50.0% LSers; Urban PRE: 57.4% LSers; Traditional PRE: 64.0% LSers). What drives this partial, yet pervasive, propensity for cognitive set in all groups? We suggest that the relative trade-offs between exploring alternative strategies and exploiting familiar ones may serve to equilibrate response styles in the absence of external biases (Ionescu, 2017). Exploring alternatives can be advantageous when a more efficient reward is discovered but it can also be risky by consuming time and resources especially when the outcome is unknown (Brosnan & Hopper, 2014; Hommel & Colzato, 2017). Thus, a balance between flexible and persistent response-styles is likely beneficial, an assertion that makes Westerners' proclivity for rule-based problem solving a worrying prospect. Future endeavors might explore other populations' relative propensities toward cognitive set in an effort to understand the balance between flexible and persistent strategy-use across humans, especially over multiple contexts.

This is the first study to explore cross-cultural differences in cognitive set. Many psychological pursuits have made conclusions based on Western participants' responses and as the spread of Western culture begins to reach even remote cultures such as the Himba (Caparos, Ahmed, et al., 2012), a process that may eventually render this type of research impossible, we argue that strengthening the current efforts toward understanding cross-cultural cognition is invaluable.

Footnotes:

[1] Many Himba exhibited great difficulty reproducing the Square1 → Square2 sequence in Training levels 1, 2 & 3 than Westerners. Thus, immediately following the initial instructions they were prompted to show the experimenter "Which came first?" then "Which came second?" Even with this extra instruction, a large portion of both Urban (55.7%) and Traditional (57.0%) Himba that began the LS-DS task did not pass the Training levels (compared to 10% of Westerners). We posit that this difficulty might be due to their lack of corrective eyewear and/or the speed of the demonstrations *(Min* = 150ms). Alternative explanations and potential implications are discussed; however, we reasoned that if increased instruction affected subjects' responses, it would only serve to concretize the LS (Crooks & McNeil, 2009).

[2] The main effect of including Training 3, although not significant [$\chi^2(1, N = 183) = 3.82$, $p = .051$], increased the predictive power of the model to 75.4% of cases. Subjects with more Training 3 trials tended to be classified as DSers. This could be interpreted as indicating that the Himba's increased experience with the Triangle could have contributed to their enhanced ability to use the shortcut. However, none of the 3 Western DSers, had abnormally increased experience with Training 3 (mean number of Training 3 trials for Westerners = 8.47), illustrating that although differences in rule familiarity could conceivably influence susceptibility to cognitive set, a causative role is unsupported.

[3] We also considered that working memory availability might influence shortcut use. In fact, Beilock and DeCaro found that, under stress, humans with less working memory availability were more likely to use the shortcut in Luchins' task than subjects with more working memory (Beilock & Decaro, 2007). In the LS-DS task, the LS requires the subject to

recall the demonstration before selecting the Triangle. Thus, lower working memory availability might result in 1) an increased number of training trials before the accuracy criterion is reached and 2) increased use of the DS, which does not require any working memory. Yet previous research has shown increased working memory availability in Urban compared with Traditional Himba (Linnell et al., 2013). Thus, if working memory played a causative role in the Himba's difficulty during the training we might have expected to see group differences in the number of training trials between Urban and Traditional Himba, but we did not.

[4] Although it has been suggested that exposure to reading/writing could promote global processing biases (Davidoff et al., 2008; Dukette & Stiles, 2001), Caparos et al. (2012) found that excluding Urban Himba who had been to school did not alter their results, Urban Himba and British subjects still exhibited a similarly global perceptual bias compared with Traditional Himbas' more local bias.

6 DISCUSSION

6.1 Summary of Findings

This research describes differences in susceptibility to cognitive set between primate species and across human cultures. Chapter 2 illustrates baboons' nearly exhaustive use of the shortcut and humans' considerable preference for the learned strategy. Chapter 3 depicts intermediate shortcut-use in chimpanzees and provides a comparative analysis of cognitive set on the LS-DS task across primate species. Chapter 4 showed that although human subjects saw the shortcut, they did not use it until their conceptualization of the problem constraints were altered. Chapter 5 found that shortcut-use varied across human cultures and presented further evidence that problem conceptualization, rather than visual perception, influences susceptibility to cognitive set in humans.

6.2 Species Differences in Shortcut-Use

Baboons and chimpanzees, but only a minority of humans, used the Triangle when it was available as a shortcut (the DS). Furthermore, chimpanzees were the only species to consistently utilize a partial shortcut, characterized by selecting the first Square(1) but then skipping Square(2) and selecting the Triangle (the SS). In the Introduction, I discussed primate behavioral flexibility across a range of tasks; however, here I will focus on the LS-DS task and the potential factors that might have contributed to the observed species' differences in shortcut-use. Keep in mind that, in the LS-DS task, the Triangle is highly salient (in PROBE trials), familiar, and has an impeccable reward history. By design, the task promotes the use of the shortcut. Thus, it is not as much surprising when the shortcut is used, as when it is not used.

6.2.1 The Frontal Cortex, Cognitive Flexibility, and Language

First, I will briefly discuss how the likely cortical underpinnings of cognitive set in primates might interact with strategy encoding. There is much evidence that regions of the frontal cortex mediate cognitive flexibility in primates (Brass & Von Cramon, 2002; Bunge et al., 2005; Bunge & Zelazo, 2006; Chrysikou et al., 2013; Gonen-Yaacovi et al., 2013; Hyafil, Summerfield, & Koechlin, 2009; Koechlin & Hyafil, 2007; Qiu et al., 2010; Roberts, 2008; Rygula, Walker, Clarke, Robbins, & Roberts, 2010; Sakai, 2008; Zelazo, 2008); note, these are by no means the only regions involved in flexible problem solving (Holmes & Cohen, 2014; Miller & Buschman, 2008; Stoet & Snyder, 2004). Within the prefrontal cortex (PFC) in particular, sensory information from the environment is integrated with previous experience, allowing a contextually appropriate response to be selected and relayed to motor areas, where the behavior is produced (Bunge & Wallis, 2008; Hoshi, 2008; Miller & Buschman, 2008; Petrides, 2008).

Importantly, the PFC has undergone extensive evolutionary changes within the primate lineage. Evidence suggests that the human neocortex is significantly larger and more convoluted than expected for our brain size and exhibits a disproportionate increase in white matter composition compared to other primates' (Rilling & Insel, 1999; Schoenemann, Sheehan, & Glotzer, 2005). In fact, although some studies suggest that relative brain size (calculated from brain to body size ratio) is similar across apes and humans (Semendeferi, Lu, Schenker, & Damasio, 2002), a recent investigation provided strong evidence for non-allometric expansions of the prefrontal cortex in apes and humans (Smaers, Gomez-Robles, Parks, & Sherwood, 2017). Furthermore, lateralization of neuroanatomical features within the PFC have been observed between monkeys, apes, and humans, which, along with cortical expansion, has been suggested to underlie species differences in cognition (Diamond & Goldman-Rakic, 1989; Hopkins, Misiura, Pope, & Latash, 2015; Rosati, 2017; Sakai, 2008; Schenker et al., 2010; Schoenemann et al., 2005; Stoet & Snyder, 2003).

Of particular note, the ventrolateral prefrontal cortex (VLPFC) has been found to be associated with various types of strategy/rule use in primates (Brass, Derrfuss, Forstmann, & von Cramon, 2005; Bunge, 2004; Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Bunge & Zelazo, 2006; Christoff & Keramatian, 2008; Crone et al., 2006; Miller & Buschman, 2008; Petrides, 2008; Roberts, 2008; Rushworth et al., 2008; Rygula et al., 2010). Additionally, in humans, the left VLPFC also houses Broca’s area (Badre & Wagner, 2007), which has long been known to be involved in language (Broca, 1861). This co-localization of rule-encoding and language within the VLPFC have led some to consider that humans’ apparently advanced use of abstractions may have developed from an enhanced ability to verbally encode rules (Sakai, 2008).

Although not necessarily a function of neuroanatomical proximity, rule encoding and verbal processing are closely linked in humans (Bahlmann, Schubotz, Mueller, Koester, & Friederici, 2009; Davidoff & Fagot, 2010; Ellis & Reingold, 2014; Jacques, 2001; Sakai, 2008; Zelazo, 2008). In fact, VLPFC activity diminishes after a rule is learned (Della-Maggiore & McIntosh, 2005; Toni, Rowe, Klass, & Passingham, 2002) and, in humans, disrupting inner speech leads to significantly longer switch costs (Emerson & Miyake, 2003). Thus, species that do not verbally encode strategies may be at a disadvantage when learning problem solving rules, especially as prescribed solutions become more complex or abstract. Indeed, I propose that differences in rule-encoding processes contribute to cognitive set by mediating how easily a rule is learned and how readily it can be replaced.

6.2.2 Species Differences in Rule Encoding

6.2.2.1 Sequential Processing

The LS is a three-step sequence, which requires subjects to identify which Square (1) came first, which Square (2) came second, and then select the Triangle. The ability to separate the Triangle from this sequence is paramount to using it as a solution by itself (i.e., the shortcut). Thus, presumably, cross-species variation in sequential processing could influence cognitive set. Specifically, processing each step of the LS sequence as an individual component [ie. (Square1) + (Square2) + (Triangle)] or even just processing the Triangle as separable from the Squares [(Square1 + Square2) + (Triangle)] might allow subjects to more readily identify the shortcut (Triangle). Conversely, processing the LS sequence as a whole construct [ie. (Square1 + Square2 + Triangle)] might render the DS more difficult to disentangle.

There are numerous similarities between humans' and nonhuman primates' abilities to process sequences (reviewed in Conway & Christiansen, 2001); however, important differences have also been reported. First, humans seem better able to conceptualize hierarchical sequences (Conway & Christiansen, 2001; Gobet et al., 2001) and, in contrast with baboons, are better able to recall structured sequences (Fagot & De Lillo, 2011). Second, there is evidence to suggest that old world monkeys might solve sequential problems by identifying and selecting each step individually, but humans and chimpanzees appeared to identify the entire sequence before reproducing it (Biro & Matsuzawa, 1999; Ohshiba, 1997). Yet, other reports suggest that both monkeys and chimpanzees are affected by a sequence in its entirety (Beran et al., 2004; Fagot & De Lillo, 2011).

There do appear to be species differences in sequential processing (Conway & Christiansen, 2001; Ghirlanda et al., 2017); however, if and to what extent these interact with the observed species differences in cognitive set on the LS-DS task is not clear. Future endeavors would benefit from comparing primate species' abilities to learn and replace a *non-sequential* learned strategy.

6.2.2.2 Rule Representations

Compared to humans, nonhuman primates require extensive training to acquire abstract strategies (Beran et al., 2004; Fagot & De Lillo, 2011; Pope et al., 2015; Stoet & Snyder, 2003). Likewise, in the present research, the amount of training that was required to learn the LS differed dramatically between nonhuman primates and humans. From the data in Chapter 2 and Chapter 3, the minimum number of training trials required for baboons and chimpanzees were 5,043 (*Max* = 20,060) and 2,784 (*Max* = 36,966), respectively; yet, 25 out of the 53 adult humans only conducted the minimum number of training trials: 24 (*Max* = 152). When analyzed, it was revealed that baboons' and chimpanzees' errors during training stemmed, not from inattention to the demonstration or from blindly reselecting recently correct squares, but from committing 'reversal' type errors, in which the correct order of Square1 and Square2 was misjudged (Chapter 3.3.1).

It is not unreasonable to suggest that humans adopted the LS more quickly due to their ability to verbally encode the LS. In fact, in humans, words are remembered more easily than nonwords but nonwords are more readily deconstructed (Ellis & Reingold, 2014). Thus, verbally identifying the rule may allow humans to remember and access abstract strategies more readily than other primates (Fagot et al., 1998; Ghirlanda et al., 2017). Indeed, perhaps baboons' and chimpanzees' inability to verbally encode the LS might have made it more difficult to learn initially but subsequently easier to replace.

6.2.2.3 Switching Strategies

Humans (Chapter 5), but not chimpanzees and baboons (Chapter 3), exhibited switch costs on the LS-DS task. In monkeys, distinct neuronal populations underlie different abstract rules within the PFC (Genovesio et al., 2005) and similarly, in humans, the location of rule-encoding activity in the PFC varies by region depending on which rule is being used (Sakai & Passingham, 2003). Likely, human rule-use includes, not only the distinct neuronal populations encoding strategies' dynamics, but also a verbal description of the rule, at least to some extent. Switching strategies might take more time for humans (resulting in switch costs) because their current strategy is verbally represented (Emerson & Miyake, 2003). If, for humans, strategy descriptions (e.g., 'choose the blue objects' or 'choose objects shaped like a boat') accompany each rule, then switching might require inhibiting and accessing a greater number of associated components of the old and new strategies, respectively. That said, given chimpanzees', but not baboons', use of the SS, it seems unlikely that they represent rules in exactly the same way. Even in simple discrimination reversal tasks, apes tended to catch on to the new rule faster than monkeys, indicating that their grasp of abstract contingencies may be better (Rumbaugh, 1971).

6.2.3 Summary

In summary, this research found that humans, but not nonhuman primates were affected by cognitive set on the LS-DS task. I posit that these species differences stem from differences in rule encoding processes, which may ultimately derive from differences in linguistic processing and/or underlying neural architecture. Indeed, the cortical machinery associated with abstract problem solving is certainly expanded in humans compared to apes, as well as in apes compared to monkeys (Rosati, 2017; Semendeferi et al., 2002) and this might go far to explain the observed species differences in susceptibility to cognitive set on the LS-DS task.

6.3 Human Variation in Shortcut-Use

Humans demonstrated substantial inter-individual differences in shortcut-use on the LS-DS task. In Chapter 2, we noted an age effect on shortcut-use. In Chapter 4, we found that conceptual but not perceptual influences seemed to impact humans' shortcut-use. And in Chapter 5, we discovered differences in shortcut-use across human cultures. Thus, clearly, there are factors that influence individuals' propensity toward cognitive set on the LS-DS task.

6.3.1 Developmental

The current research noted that children utilized the DS in PROBE trials significantly more often than adolescents or adults (Chapter 2; Pope et al., 2015). However, reanalysis using the 'true' measure of DS-use, in which BASE DS-use is subtracted from PROBE DS-use, this trend (children = 18.5% DSers, adolescents = 4% DSers, and adults = 9.6% DSers) was no longer significant (See Chapter 3 Footnote 2), suggesting that erroneous DS-use may have been driving the age-effects. It is also possible that this trend was driven by the younger children (3 of the 5 'true' DSers within the 7-10 age group were 7 years old) and that children under 7 years old might be even better able to utilize the DS; however, further investigation is required.

Despite somewhat mixed findings regarding human development and cognitive flexibility, it has been posited that a juvenile period of plasticity may be favorable (even selected for) because it might encourage, or simply allow, a chance for more exploratory learning (Gopnik et al., 2015; Griffin, 2016; Ionescu, 2017; Thompson-Schill et al., 2009). In fact, Ionescu (2017) posited that when learning a behavioral strategy, humans engage in an initial exploratory period, followed by a period of stability, followed by another period of flexibility. It is plausible that this so-called 'variability-stability-flexibility' pattern underlies human differences, even beyond developmental differences, in one's propensity to adopt the shortcut on the LS-DS task.

6.3.2 Working memory

In Chapter 2, we suggested that working memory availability might influence susceptibility to cognitive set. Specifically, we posited that baboons' and children's enhanced propensities to use the shortcut might stem from their limited working memory availability compared to adults (Fagot & De Lillo, 2011; Miles et al., 1996; Thomason et al., 2009), such that the LS is inherently more difficult (as it requires remembering the locations of Square1 and Square2) making the DS more attractive by comparison (see Beilock & Decaro, 2007). However, the data from Chapter 5 do not support this contention; although Traditional and Urban Himba differ in their working memory capacities (Linnell et al., 2013), we did not find a consistent difference in their DS-use. Furthermore, preliminary findings from a modified LS-DS task hint that, when the LS working memory requirements are alleviated, humans may actually utilize the DS *more*, while capuchins and rhesus macaques exhibit DS-use similar to that of baboons (Watzek & Pope, Unpublished). Thus, although certainly not concrete, the available data do not suggest that higher working memory availability – in and of itself – promotes inflexibility on the LS-DS task. However, future investigations might also consider increasing cognitive load during testing to examine this possibility further.

6.3.3 The Influence of Problem Conceptualization on Cognitive Set

Chapters 4 and 5 made use of interventions, midway through testing, aimed at altering subjects' conceptualization of the problem. In Chapter 4, we showed half the participants a video that demonstrated the DS (*Min* = 8 times) and in Chapter 5, we told them "Don't be afraid to try new things." Increased shortcut-use was observed following both manipulations, indicating that understanding that the shortcut was a viable strategy, effectively enabled subjects to use it. This is not surprising. Essentially, Chapter 4 taught subjects how to use the shortcut and Chapter 5 clearly implies the possibility of multiple solutions. However we observed two outcomes that were unexpected.

First, after viewing the DS demonstration (Chapter 4), five times as many subjects were classified as DSers; however 31% *still* did not use the shortcut. Simply put, a subset of subjects watched a video that showed the task being solved via the shortcut – and still did not use it, thereby demonstrating the strength of set in some individuals. Seemingly, subjects were so convinced that they had already identified the solution, they did not bother attending to the demonstration. This is frighteningly reminiscent of confirmation bias, wherein incoming information is inadvertently modulated to fit preconceived conceptions (Bilalić et al., 2008; Doll, Hutchison, & Frank, 2011). Indeed, this lack of exploratory behavior goes far towards explaining why certain subjects are able to break cognitive set.

The second unexpected finding is reported in Chapter 5. Following the "don't be afraid to try new things" prompt, nearly nine times more Westerners used the shortcut; yet, Himba participants' shortcut-use did not increase. One possibility is that something was lost in the translation of the prompt. However, this is unlikely given that Urban Himba's DS-use significantly *decreased* in POST trials. It seems that, because many Himba were already using the shortcut, adhering to the "Don't be afraid to try new things" prompt meant *Use a different strategy*, resulting in decreased DS-use. Furthermore, we observed a similar response in Chapter 4 when, after Control subjects viewed the video demonstration of the LS, the number of DSers *dropped* from 13.8% to 6.9%. Similarly, it appears as if they were conforming to the demonstration, or what they thought they *should* do.

6.3.4 Summary

Humans' were better able to break away from their cognitive set when their conceptual understanding of the task changed. However, in some cases, not even a demonstration of the DS was able to alter subjects' conceptualization of the LS as *the* solution. I suggest that cognitive set is, to some extent, a byproduct of codified rule-use such that strategic representations, once formed, might be difficult to replace. That said, there is substantial inter-individual variability in both initial susceptibility to cognitive set and the propensity to break it in humans.

6.4 Broader Implications

Altering subjects' conceptualization of the problem effectively alters their ability to break cognitive set; however, what pre-existing conceptualizations might allow some subjects to be less affected by cognitive set to begin with is unclear. The following sections present several speculative hypotheses on this topic, as well as thoughts for future directions.

6.4.1 *Cultural Influences*

Throughout this research, remarkably few Western subjects were classified as DSers, prior to intervention (See Table 6.3-1). However, in stark contrast, 38% of the remote Himba adults adopted the shortcut within their first 48 trials. As Chapters 4 and 5 showed, perceptual influences do not appear to impact shortcut-use on the LS-DS. Thus, I will not discuss the known cultural differences in perceptual processing. Instead, I will focus on cultural differences that might impact problem conceptualization.

Table 6.1 The proportions of human populations classified as DSers.

	Nationality	**N**	**%DSers**
Chapter 2/3:	American	53	7.5%
Pilot:	French	14	7.1%
Chapter 4:	American	58	13.8%
Chapter 5:	American	54	5.6%
	Namibian	129	38.0%

Westerners and Himba participants differ in many ways, including social structure, physical environment, language, and educational background. Here, we will focus on the potential influence of language and educational background. Note that, although group differences in genetic predispositions biasing problem-solving approach toward either persistence or flexibility are certainly possible (Hommel & Colzato, 2017), I find this an unlikely explanation given the extremely diverse population of Western students sampled: out of 904 students active on Georgia State University's SONA system (as of November 2nd, 2017), 17% identify as Asian, 50% identify as Black/African American, 23% identify as White/Caucasian, and 10% identified as more than one race.

6.4.2 The Impact of Language on Cognitive Set

As discussed briefly in Section 6.2.1, for humans, language provides the scaffolding upon which abstract constructs are compiled into problem-solving strategies (Emerson & Miyake, 2003; Lipton & Eichenbaum, 2008). Humans are capable of performing novel actions, even a complex series of actions, in the complete absence of demonstration and without resorting to trial and error, simply by following a set of verbal or written instructions (reviewed by Stoet & Snyder, 2008). This cognitive hack allows a large number of people to benefit from the knowledge of a few (e.g., textbooks). However, the close association between language and rule use means that differences in linguistic encoding might easily interact with strategy-use (Jacques, 2001).

Recall the assertion that the observed differences in susceptibility to cognitive set across primate species might stem from differences in rule-encoding, potentially underlain by the ability to verbally encode abstract rules (discussed in Section 6.2.2.2). The Himba do not have a direct translation for the word 'shapes.' Thus, their ability to encode the LS might have suffered to some extent and conceivably influenced their susceptibility to cognitive set. Indeed, Himba participants had much more difficulty passing the training levels and the ones that did, required significantly more training trials ($M = 79.7$) than Western participants ($M = 39$; Section 5.3.5). That being said, all Himba participants were monolingual in Otjihimba; yet we observed substantial inter-individual variation in susceptibility to cognitive set and there was no interaction between the number of training trials and DS-use in humans. Furthermore, consider that *all* of the human populations tested required thousands fewer training trials than the nonhuman subjects and the *majority* of each group exhibited a preference for the LS.

Thus, it seems reasonable that humans possess roughly equivalent abilities to encode rules; although, certainly, formally educated participants might be more practiced in this regard. Indeed, enhanced rule encoding, perhaps stemming from linguistic processes, might promote the initial development of cognitive set in humans but this appears to have played a secondary role in the observed inter-individual differences in humans' abilities to break cognitive set. Specifically, as evidenced in Chapters 4 and 5, subjects' conceptualization of the task appears to greatly influence their abilities to use the shortcut.

6.4.3 Potential Impact of Educational Background

Western education relies, in large part, on rote memorization of prescribed solution strategies to boost efficiency and long-term retention (Fehr, 1953; Henderson & Pingry, 1953; Schwartz, Bransford, & Sears, 2005). However, this approach might well be ineffective (Rohrer & Taylor, 2006) and/or promote inflexibility (Sweller, 1983). I have discussed this possibility throughout the text (Sections: 1.5.1; 1.6.4; 2.1; 4.1; 4.4; 5.1; 5.4); however, the current research only *hints* at the possibility that educational practices might contribute to set. That being said, I would like to briefly delve into the potential implications and practicalities.

First, it is also possible that exposure to formal education promotes the interpretation of the experimental setup as a test of sorts, such that a fear of failure prevents exploration (Luchins, 1942). This goes a long way to explain why the "Don't be afraid to try new things" prompt effectively reduced cognitive set in Western but not Himba participants (Section 5.3.2). An interesting manipulation might be to manipulate the reward values of correct (increase in reward) and incorrect (neutral vs decrease in reward) responses.

It is important to note in this discussion that, undoubtedly, mechanized rule-use has its benefits: it lightens cognitive load and allows many different problems to be solved via the same strategy (Christoff & Keramatian, 2008; Sweller, 1980; Sweller et al., 1982). Furthermore, a recent meta-analysis reported that one year of education corresponds to an approximate increase of 1-5 IQ points (Ritchie & Tucker-Drob, 2017). It would be impractical to completely forego problem sets as educational tools to learn specific problem solving methods; however, a seemingly common frustration among educators is that children do not know how to "learn for themselves" (Schwartz et al., 2005). How do we balance "Exploit!" and "Explore!" in the learning environment? A better understanding of how, exactly, cognitive flexibility is influenced by educational styles and to what extent this might intersect with personality traits would be a worthy pursuit for future endeavors. In particular, comparing Western and non-Western children who either do or do not partake in formal education (e.g., public, private), in a longitudinal study, would go far to elucidate the impacts of education on cognitive flexibility.

6.5 Concluding Remarks

Strategy selection can, and arguably should (Brosnan & Hopper, 2014), be biased by situational factors (Hommel & Colzato, 2017; Kolodny et al., 2015). Effective strategies should be maintained within both individual and group repertoires such that, when another solution method becomes necessary or available, it can be capitalized upon (Reader, 2016). On a proximate level, we might expect individuals with better faculty over explorative/exploitative strategies to exhibit higher fitness, especially in dynamic environments (Carr et al., 2016; Holmes & Cohen, 2014; Reader, 2003).

Humans exhibit a remarkable proclivity for invention; yet time and again, engage in mechanized problem solving (Adamson, 1952; Bilalić et al., 2008; Chrysikou & Weisberg, 2005;

Lemaire & Leclere, 2014; Luchins, 1942; Luchins & Luchins, 1950). How can humans' sporadic yet conclusively innovative behavior be reconciled with their known predilection for habit? I submit that explorative/exploitative tendencies exist in equilibrium within individuals such that the relative risks and rewards are optimized. Furthermore, I would speculatively propose that typical Western educational practices may impair human cognitive flexibility, by heavily rewarding habit-based solutions.

Although learned responses or habit-based behavior is not always detrimental, defaulting toward exploitative rather than explorative behavior doubtlessly contributes to inefficient responses. In this research we observed striking species differences in susceptibility to cognitive set between baboons, chimpanzees, and humans. This research also found evidence *against* the contention that perceptual influences impact shortcut-use on the LS-DS task. However, problem conceptualization was shown to impact subjects' abilities to break cognitive set. Furthermore, this is the first research to document cross-cultural differences in cognitive set in humans. I speculatively suggest that formal education might contribute to problem solving inflexibility; however future research is necessary to substantiate this claim.

REFERENCES

Abel, T. M., & Hsu, F. L. K. (1949). Some aspects of personality of Chinese as revealed by the Rorschach Test. *Rorschach Research Exchange and Journal of Projective Techniques, 13*(3), 285-301.

Adamson, R. E. (1952). Functional fixedness as related to problem solving: A repetition of three experiments. *Journal of Experimental Psychology, 44*(4), 288-291.

Aftanas, M. S., & Koppenaal, R. J. (1962). Effects of instructional problems and jar position variation on the water-jar einstellung test. *Psychological Reports, 10,* 359-362.

Amici, F., Aureli, F., & Call, J. (2008). Fission-fusion dynamics, behavioral flexibility, and inhibitory control in primates. *Curr Biol, 18*(18), 1415-1419. doi:10.1016/j.cub.2008.08.020

Ardiale, E., & Lemaire, P. (2012). Within-item strategy switching: An age of comparative study in adults. *Psychology and Aging, 27*(4), 1138-1151. doi:10.1037/a0027772

Arnett, J. J. (2008). The neglected 95%: why American psychology needs to become less American. *Am Psychol, 63*(7), 602-614. doi:10.1037/0003-066X.63.7.602

Arrington, C. M., & Logan, G. D. (2004). The cost of voluntary task switch. *Psychological Science, 15*(9), 610-615.

Avila, J. (2013). *The effects of acculturation on neuropsychological test performance in English-speaking and Farsi-speaking Iranian individuals.* (Master of Arts in Psychology), California State University, Northridge.

Badre, D. (2008). Ventrolateral Prefrontal Cortex and Controlling Memory to Inform Action. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behavior* (pp. 365-389). New York: Oxford University Press.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901. doi:10.1016/j.neuropsychologia.2007.06.015

Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., & Friederici, A. D. (2009). Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Res, 1298,* 161-170. doi:10.1016/j.brainres.2009.08.017

Barth, J., & Call, J. (2006). Tracking the displacement of objects: a series of tasks with great apes (Pan troglodytes, Pan paniscus, Gorilla gorilla, and Pongo pygmaeus) and young children (Homo sapiens). *J Exp Psychol Anim Behav Process, 32*(3), 239-252. doi:10.1037/0097-7403.32.3.239

Beilock, S. L., & Decaro, M. S. (2007). From poor performance to success under stress: working memory, strategy selection, and mathematical problem solving under pressure. *J Exp Psychol Learn Mem Cogn, 33*(6), 983-998. doi:10.1037/0278-7393.33.6.983

Beran, M. J., Pate, J. L., Washburn, D. A., & Rumbaugh, D. M. (2004). Sequential responding and planning in chimpanzees (Pan troglodytes) and rhesus macaques (Macaca mulatta). *J Exp Psychol Anim Behav Process, 30*(3), 203-212. doi:10.1037/0097-7403.30.3.203

Beran, M. J., Rossettie, M. S., & Parrish, A. E. (2016). Trading up: chimpanzees (Pan troglodytes) show self-control through their exchange behavior. *Anim Cogn, 19*(1), 109-121. doi:10.1007/s10071-015-0916-7

Beran, M. J., & Washburn, D. A. (2002). Chimpanzee responding during matching to sample: control by exclusion. *Journal of the experimental analysis of behavior, 78*(3), 497-508.

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.

Berg, E. A. (1948). A Simple Objective Technique for Measuring Flexibility in Thinking. *The Journal of General Psychology, 39*(1), 15-22. doi:10.1080/00221309.1948.9918159

Berl, R. E., & Hewlett, B. S. (2015). Cultural variation in the use of overimitation by the Aka and Ngandu of the Congo Basin. *PLoS One, 10*(3), e0120180. doi:10.1371/journal.pone.0120180

Besnard, D., & Cacitti, L. (2005). Interface changes causing accidents. An empirical study of negative transfer. *International Journal of Human-Computer Studies, 62*(1), 105-125.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition, 108*(3), 652-661.

Biro, D., & Matsuzawa, T. (1999). Numerical ordering in a chimpanzee (Pan troglodytes): Planning, executing, and monitoring. *Journal of Comparative Psychology, 113*(2), 178.

Bonté, E., Flemming, T., & Fagot, J. . (2011). Executive control of perceptual features and abstract relations by baboons (Papio papio). *Behavioral Brain Research, 222*(1), 176-182.

Bonte, E., Kemp, C., & Fagot, J. (2014). Age effects on transfer index performance and executive control in baboons (Papio papio). *Front Psychol, 5*, 188. doi:10.3389/fpsyg.2014.00188

Brass, M., Derrfuss, J., Forstmann, B., & von Cramon, D. Y. (2005). The role of the inferior frontal junction area in cognitive control. *Trends Cogn Sci, 9*(7), 314-316. doi:10.1016/j.tics.2005.05.001

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2008). The role of the Posterior frontolateral cortex in task-related control` *Neuroscience of Rule-Guided Behavior* (pp. 177-195). New York: Oxford University Press.

Brass, M., & Von Cramon, D. Y. (2002). The role of the frontal cortex in task preparation. *Cerebral Cortex, 12*(9), 908-914.

Broca, P. (1861). Remarques sur la sige de la faculté du langage articulé, suivi d'une observation d'aphémie. . *Bulletin de la Societé Anatomique de Paris, 2*, 300-357.

Brosnan, S. F., & Hopper, L. M. (2014). Psychological limits on animal innovation. *Animal Behaviour, 92*, 325-332. doi:10.1016/j.anbehav.2014.02.026

Bunge, S. A. (2004). How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cognitive, Affective, and Behavioral Neuroscience, 4*(4), 564-579.

Bunge, S. A., Kahn, I., Wallis, J. D., Miller, E. K., & Wagner, A. D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *J Neurophysiol, 90*, 3419-3428.

Bunge, S. A., & Wallis, J. D. (2008). Introduction. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behavior* (pp. xiii-xxiii). New York: Oxford University Press.

Bunge, S. A., Wallis, J. D., Parker, A., Brass, M., Crone, E. A., Hoshi, E., & Sakai, K. (2005). Neural circuitry underlying rule use in humans and nonhuman primates. *J Neurosci, 25*(45), 10347-10350. doi:10.1523/JNEUROSCI.2937-05.2005

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science, 15*(3), 118-121.

Buttelmann, F., & Karbach, J. (2017). Development and Plasticity of Cognitive Flexibility in Early and Middle Childhood. *Front Psychol, 8*, 1040. doi:10.3389/fpsyg.2017.01040

Camerer, C. F., & Weber, M. (1991). Recent developments in modeling preferences: Uncertainty and ambiguity. *Manusckipte aus den Instituten fur Betriebswirtschaftslehre der Universitat Kiel, 275*.

Caparos, S., Ahmed, L., Bremner, A. J., de Fockert, J. W., Linnell, K. J., & Davidoff, J. (2012). Exposure to an urban environment alters the local bias of a remote culture. *Cognition, 122*(1), 80-85. doi:10.1016/j.cognition.2011.08.013

Caparos, S., Linnell, K. J., Bremner, A. J., de Fockert, J. W., & Davidoff, J. (2012). Does local/global perceptual bias tell us anything about local/global selective attention? *Visual Cognition, 20*(9), 1016-1020.

Carr, K., Kendal, R. L., & Flynn, E. G. (2016). Eureka!: What Is Innovation, How Does It Develop, and Who Does It? *Child Dev, 87*(5), 1505-1519. doi:10.1111/cdev.12549

Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Cart Sorting test. *Journal of Clinical and Experimental Neuropsychology, 8*(3), 219-228.

Chevalier, N., Blaye, A., Dufau, S., & Lucenet, J. (2010). What visual information do children and adults consider while switching between tasks? Eye-tracking investigation of cognitive flexibility development. *Dev Psychol, 46*(4), 955-972. doi:10.1037/a0019674

Chiu, L. (1972). A Cross-Cultural Comparison of Cognitive Styles in Chinese and American Children. *International Journal of Psychology, 7*(4), 235-242. doi:10.1080/00207597208246604

Christoff, K., & Keramatian, K. (2008). Abstraction of Mental Representations: Theoretical Considerations and Neuroscientific Evidence. In S. A. Bunge & J. D. Wallis (Eds.), *The Neuroscience of Rule-Guided Behavior* (pp. 107-126). New York: Oxford University Press.

Chrysikou, E. G., Hamilton, R. H., Coslett, H. B., Datta, A., Bikson, M., & Thompson-Schill, S. L. (2013). Noninvasive transcranial direct current stimulation over the left prefrontal cortex facilitates cognitive flexibility in tool use. *Cogn Neurosci, 4*(2), 81-89. doi:10.1080/17588928.2013.768221

Chrysikou, E. G., & Weisberg, R. W. (2005). Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task. *J Exp Psychol Learn Mem Cogn, 31*(5), 1134-1148. doi:10.1037/0278-7393.31.5.1134

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proc Natl Acad Sci U S A, 102*(35), 12629-12633. doi:10.1073/pnas.0506162102

Clegg, J. M., & Legare, C. H. (2016). A cross-cultural comparison of children's imitative flexibility. *Dev Psychol, 52*(9), 1435-1444. doi:10.1037/dev0000131

Coffey, D. M., Marmol, L., Schock, L., & Adams, W. (2005). The influence of acculturation on the Wisconsin Card Sorting Test by Mexican Americans. *Arch Clin Neuropsychol, 20*(6), 795-803. doi:10.1016/j.acn.2005.04.009

Cole, P., Duncan, L. G., & Blaye, A. (2014). Cognitive flexibility predicts early reading skills. *Front Psychol, 5*, 565. doi:10.3389/fpsyg.2014.00565

Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends Cogn Sci, 5*(12), 539-546.

Cools, R. (2008). Dopaminergic Modulation of flexible cognitive control: the role of the striatum. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behaviors* (pp. 313-334). New York: Oxford University Press.

Corballis, M. C. (2010). Mirror neurons and the evolution of language. *Brain Lang, 112*(1), 25-35. doi:10.1016/j.bandl.2009.02.002

Crone, E. A., Donohue, S. E., Honomichl, R., Wendelken, C., & Bunge, S. A. (2006). Brain regions mediating flexible rule use during development. *J Neurosci, 26*(43), 11239-11247. doi:10.1523/JNEUROSCI.2165-06.2006

Crooks, N. M., & McNeil, N. M. (2009). Increased practice with 'set' problems hinders performance on the water jar task. *Preceedings of the 31st annual conference of the Cognitive Science Society, Cognitive Science Society*, 643-648.

Cunningham, J. D. (1965). Einstellung rigidity in children. *Journal of Experimental Child Psychology, 2*(237-247).

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature, 398*, 203-204.

Davidoff, J., & Fagot, J. (2010). Cross-species Assessment of the Linguistic Origins of Color Categories. *Comparative Cognition & Behavior Reviews, 5*, 100-116. doi:10.3819/ccbr.2010.50005

Davidoff, J., Fonteneau, E., & Fagot, J. (2008). Local and global processing: observations from a remote culture. *Cognition, 108*(3), 702-709. doi:10.1016/j.cognition.2008.06.004

Davis, S. J. (2017). *The Context of Behavioural Flexibility in Chimpanzees (Pan troglodytes): Implications for the Evolution of Cumulative Culture.* (PhD in Psychology), University of St. Andrews.

Davis, S. J., Vale, G. L., Schapiro, S. J., Lambeth, S. P., & Whiten, A. (2016). Foundations of cumulative culture in apes: improved foraging efficiency through relinquishing and combining witnessed behaviours in chimpanzees (Pan troglodytes). *Sci Rep, 6*, 35953. doi:10.1038/srep35953

de Fockert, J. W., Caparos, S., Linnell, K. J., & Davidoff, J. (2011). Reduced distractability in a remote culture. *PLoS One, 6*(10), e26337. doi:10.1371/

de Fockert, J. W., Davidoff, J., Fagot, J., Parron, C., & Goldstein, J. (2007). More accurate size contrast judgments in the Ebbinghaus Illusion by a remote culture. *J Exp Psychol Hum Percept Perform, 33*(3), 738-742. doi:10.1037/0096-1523.33.3.738

de Waal, F. B., & Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends Cogn Sci, 14*(5), 201-207. doi:10.1016/j.tics.2010.03.003

Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science, 335*(6072), 1114-1118. doi:10.1126/science.1213969

Defeyter, M. A., & German, T. P. (2003). Acquiring an understanding of design: evidence from children's insight problem solving. *Cognition, 89*(2), 133-155. doi:10.1016/s0010-0277(03)00098-2

Della-Maggiore, V., & McIntosh, A. R. (2005). Time course of changes in brain activity and functional connectivity associated with long-term adaptation to a rotational transformation. *J Neurophysiol, 93*(4), 2254-2262. doi:10.1152/jn.00984.2004

Deruelle, C., & Fagot, J. (1998). Visual search for global/local stimulus features in humans and baboons. *Psychonomic Bulletin & Review, 5*(3), 476-481.

Diamond, A., & Goldman-Rakic, P. S. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: evidence for dependence on dorsolateral prefrontal cortex. *Exp Brain Res, 74*, 24-40.

Diamond, A., & Kirkham, N. (2005). Not quite as grown-up as we like to think: parallels between cognition in childhood and adulthood. *Psychol Sci, 16*(4), 291-297. doi:10.1111/j.0956-7976.2005.01530.x

Dolins, F. L., Schweller, K., & Milne, S. (2017). Technology advancing the study of animal cognition: using virtual reality to present virtually simulated environments to investigate nonhuman primate spatial cognition. *Current Zoology, 63*(1), 97-108. doi:10.1093/cz/zow121

Doll, B. B., Hutchison, K. E., & Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J Neurosci, 31*(16), 6188-6198. doi:10.1523/JNEUROSCI.6486-10.2011

Dukette, D., & Stiles, J. (2001). The effects of stimulus density on children's analysis of hierarchical patterns. *Developmental Science, 4*(2), 233-251.

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological Monographs, 58*(5), i-113. doi:10.1037/h0093599

Ellis, J. J., & Reingold, E. M. (2014). The Einstellung effect in anagram problem solving: evidence from eye movements. *Front Psychol, 5*, 679. doi:10.3389/fpsyg.2014.00679

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48*, 148-168.

Fagot, J., & Bonte, E. (2010). Automated testing of cognitive performance in monkeys: use of a battery of computerized test systems by a troop of semi-free-ranging baboons (Papio papio). *Behav Res Methods, 42*(2), 507-516. doi:10.3758/BRM.42.2.507

Fagot, J., & De Lillo, C. (2011). A comparative study of working memory: immediate serial spatial recall in baboons (Papio papio) and humans. *Neuropsychologia, 49*(14), 3870-3880. doi:10.1016/j.neuropsychologia.2011.10.003

Fagot, J., & Deruelle, C. (1997). Processing of global and local visual information and hemispheric specialization in Humans (*Homo sapiens*) and Baboons (*Papio papio*). *Journal of experimental Psychology: Human Perception and Performance, 23*(2), 429-442.

Fagot, J., Goldstein, J., Davidoff, J., & Pickering, A. (2006). Cross-species differences in color categorization. *Psychonomic Bulletin & Review, 13*(2), 275-280.

Fagot, J., Kruschke, J. K., Depy, D., & Vauclair, J. (1998). Associative learning in baboons (Papio papio) and humans (Homo sapiens): species differences in learned attention to visual features. *Animal Cognition, 1*, 123-133.

Fagot, J., & Paleressompoulle, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behav Res Methods, 41*(2), 396-404. doi:10.3758/BRM.41.2.396

Fagot, J., & Tomonaga, M. (1999). Global and Local Processing in Humans (Homo sapiens) and Chimpanzees (Pan troglodytes): Use of a Visual Search Task With Compound Stimuli. *Journal of Comparative and Physiological Pyschology, 113*(1), 3-12.

Fehr, H. F. (1953). The learning of mathematics: Its theory and practice. *National Council of Teachers of Mathematics, 21*.

Fragaszy, D. M., Kennedy, E., Murnane, A., Menzel, C., Brewer, G., Johnson-Pynn, J., & Hopkins, W. (2009). Navigating two-dimensional mazes: chimpanzees (Pan troglodytes) and capuchins (Cebus apella sp.) profit from experience differently. *Anim Cogn, 12*(3), 491-504. doi:10.1007/s10071-008-0210-z

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: evidence from Piraha language and cognition. *Cognition, 108*(3), 819-824. doi:10.1016/j.cognition.2008.04.007

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J Exp Psychol Gen, 137*(2), 201-225. doi:10.1037/0096-3445.137.2.201

Gallup, G. G., Jr. (1970). Chimpanzees: Self-Recognition. *Science, 167*(3914), 86-87.

Gallup, G. G., Jr. (1976). Absence of self-recognition in a monkey (*Macaca fascicularis*) following prolonged exposure to a mirror. *Dev Psychol, 10*(3), 281-284.

Genovesio, A., Brasted, P. J., Mitz, A. R., & Wise, S. P. (2005). Prefrontal cortex activity related to abstract response strategies. *Neuron, 47*(2), 307-320. doi:10.1016/j.neuron.2005.06.006

Genovesio, A., & Wise, S. P. (2008). The Neurophysiology of Abstract Response Strategies. In S. A. Bunge & J. D. Wallis (Eds.), *The Neuroscience of Rule-Guided Behavior* (pp. 81-105). New York: Oxford University Press.

German, T. P., & Barrett, H. C. (2005). Functional Fixedness in a Technologically Sparse Culture. *American Psychological Society, 16*(1), 1-4.

German, T. P., & Defeyter, M. A. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review, 7*(4), 707-712.

Ghirlanda, S., Lind, J., & Enquist, M. (2017). Memory for stimulus sequences: a divide between humans and other animals? *R Soc Open Sci, 4*(6), 161011. doi:10.1098/rsos.161011

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends Cogn Sci, 5*(6), 236-243.

Gonen-Yaacovi, G., de Souza, L. C., Levy, R., Urbanski, M., Josse, G., & Volle, E. (2013). Rostral and caudal prefrontal contribution to creativity: a meta-analysis of functional imaging data. *Front Hum Neurosci, 7*, 465. doi:10.3389/fnhum.2013.00465

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When Younger Learners Can Be Better (or at Least More Open-Minded) Than Older Ones. *Current Directions in Psychological Science, 24*(2), 87-92. doi:10.1177/0963721414556653

Griffin, A. S. (2016). Innovativeness as an emergent property: a new alignment of comparative and experimental research on animal innovation. *Philos Trans R Soc Lond B Biol Sci, 371*(1690). doi:10.1098/rstb.2015.0544

Gruber, T. (2016). Great Apes Do Not Learn Novel Tool Use Easily: Conservatism, Functional Fixedness, or Cultural Influence? *International Journal of Primatology, 37*(2), 296-316. doi:10.1007/s10764-016-9902-4

Gruber, T., Muller, M. N., Reynolds, V., Wrangham, R., & Zuberbuhler, K. (2011). Community-specific evaluation of tool affordances in wild chimpanzees. *Sci Rep, 1*, 128. doi:10.1038/srep00128

Hanus, D., Mendes, N., Tennie, C., & Call, J. (2011). Comparing the performances of apes (Gorilla gorilla, Pan troglodytes, Pongo pygmaeus) and human children (Homo sapiens) in the floating peanut task. *PLoS One, 6*(6), e19555. doi:10.1371/journal.pone.0019555

Harlow, H. (1949). The formation of learning sets. *Pscyhological Review, 56*(1), 51-65.

Haun, D. B., & Call, J. (2008). Imitation recognition in great apes. *Curr Biol, 18*(7), R288-290. doi:10.1016/j.cub.2008.02.031

Henderson, K. B., & Pingry, R. E. (1953). Problem-Solving in Mathematics. In H. F. Fehr (Ed.), *The Learning of Mathematics: Its Theory and Practice* (Vol. 21, pp. 228-270). Reston, VA: National Council of Teachers of Mathematics.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behav Brain Sci, 33*(2-3), 61-83; discussion 83-135. doi:10.1017/S0140525X0999152X

Heyes, C. M. (1993). Imitation, culture and cognition. *Anim Behav, 46*, 999-1010.

Hobaiter, C., & Byrne, R. W. (2010). Able-bodied wild chimpanzees imitate a motor procedure used by a disabled individual to overcome handicap. *PLoS One, 5*(8), e11959. doi:10.1371/journal.pone.0011959

Hockings, K. J., McLennan, M. R., Carvalho, S., Ancrenaz, M., Bobe, R., Byrne, R. W., & Dunbar, R. I. (2015). Apes in the Anthropocene: flexibility and survival. *Trends in Ecology & Evolution, 30*(4), 215-222.

Holmes, P., & Cohen, J. D. (2014). Optimality and some of its discontents: successes and shortcomings of existing models for binary decisions. *Top Cogn Sci, 6*(2), 258-278. doi:10.1111/tops.12084

Hommel, B., & Colzato, L. S. (2017). The social transmission of metacontrol policies: Mechanisms underlying the interpersonal transfer of persistence and flexibility. *Neurosci Biobehav Rev*. doi:10.1016/j.neubiorev.2017.01.009

Hopkins, W. D. (1997). Hemispheric specialization for local and global processing of hierarchical visual stimuli in chimpanzees "Pan troglodytes#. *Neuropsychologia, 35*(3), 343-348.

Hopkins, W. D., Misiura, M., Pope, S. M., & Latash, E. M. (2015). Behavioral and brain asymmetries in primates: a preliminary evaluation of two evolutionary hypotheses. *Ann N Y Acad Sci, 1359*, 65-83. doi:10.1111/nyas.12936

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*(1), 27-31.

Hopper, L. M., Kurtycz, L. M., Ross, S. R., & Bonnie, K. E. (2015). Captive chimpanzee foraging in a social setting: a test of problem solving, flexibility, and spatial discounting. *PeerJ, 3*, e833. doi:10.7717/peerj.833

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Anim Cogn, 8*(3), 164-181. doi:10.1007/s10071-004-0239-6

Hoshi, E. (2008). Differential involvement of the Prefrontal, Premotor, and Primary motor cortices in rule-based motor behavior *Neuroscience of Rule-Guided Behavior* (pp. 159-175). New York: Oxford University Press.

Hovers, E. (2012). Invention, Reinvention and Innovation *Origins of Human Innovation and Creativity* (pp. 51-68).

Hrubesch, C., Preuschoft, S., & van Schaik, C. (2009). Skill mastery inhibits adoption of observed alternative solutions among chimpanzees (Pan troglodytes). *Anim Cogn, 12*(2), 209-216. doi:10.1007/s10071-008-0183-y

Huber, L., Range, F., Voelkl, B., Szucsich, A., Viranyi, Z., & Miklosi, A. (2009). The evolution of imitation: what do the capacities of non-human animals tell us about the mechanisms of imitation? *Philos Trans R Soc Lond B Biol Sci, 364*(1528), 2299-2309. doi:10.1098/rstb.2009.0060

Huguet, P., Barbet, I., Belletier, C., Monteil, J. M., & Fagot, J. (2014). Cognitive control under social influence in baboons. *J Exp Psychol Gen, 143*(6), 2067-2073. doi:10.1037/xge0000026

Hyafil, A., Summerfield, C., & Koechlin, E. (2009). Two mechanisms for task switching in the prefrontal cortex. *J Neurosci, 29*(16), 5135-5142. doi:10.1523/JNEUROSCI.2828-08.2009

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Curr Biol, 17*(23), R1004-1005. doi:10.1016/j.cub.2007.10.027

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology, 30*(2), 190-200. doi:10.1016/j.newideapsych.2011.11.001

Ionescu, T. (2017). The variability-stability-flexibility pattern: A possible key to understanding the flexibility of the human mind. *Review of General Psychology, 21*(2), 123-131. doi:10.1037/gpr0000110

Jacques, S. (2001). *The roles of labelling and abstraction in the development of cognitive flexibility.* (PhD), University of Toronto, National Library of Canada.

Jahoda, G. (2016). Seventy years of social psychology: A cultural and personal critique. *Journal of Social and Political Psychology, 4*(1), 364-380. doi:10.5964/jspp.v4i1.621

Janzen, H. L., Maguire, T. O., & Boersma, F. J. (1976). *A developmental analysis of set patterns in children: A normative study.* Paper presented at the Annual Meeting of the American Educational Research Association, San Fransisco, CA.

Ji, L. J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology, 78*, 943-955.

Ji, L. J., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *J Pers Soc Psychol, 87*(1), 57-65. doi:10.1037/0022-3514.87.1.57

Jordan, P. L., & Morton, J. B. (2012). Perseveration and the status of 3-year-olds' knowledge in a card-sorting task: evidence from studies involving congruent flankers. *J Exp Child Psychol, 111*(1), 52-64. doi:10.1016/j.jecp.2011.06.011

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*(5), 449-467. doi:10.1111/1467-7687.00300

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: a cultural look at new look. *Psychol Sci, 14*, 201-206.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition, 29*(7), 1000-1009.

Koechlin, E., & Hyafil, A. (2007). Anterior prefrontal function and the limits of human decision-making. *Science, 318*(5850), 594-598. doi:10.1126/science.1142995

Kolodny, O., Edelman, S., & Lotem, A. (2015). Evolved to adapt: A computational approach to animal innovation and creativity. *Current Zoology, 61*(2), 350-367.

Kuhnen, U., & Oyserman, D. (2002). Thinking about the self influences thinking in general: cognitive consequences of salient self-concept. *Journal of Experimental Social Psychology, 38*, 492-499.

Laland, K. N. (2004). Social learning strategies. *Learning & Behavior, 32*(1), 4-14.

Langer, E. J., & Lois, G. I. (1979). When practice makes imperfect: debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37*(11), 2014.

Legare, C. H., & Nielsen, M. (2015). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends Cogn Sci, 19*(11), 688-699. doi:10.1016/j.tics.2015.08.005

Lehner, S. R., Burkart, J. M., & Schaik, C. P. (2011). Can captive orangutans (Pongo pygmaeus abelii) be coaxed into cumulative build-up of techniques? *J Comp Psychol, 125*(4), 446-455. doi:10.1037/a0024413

Lemaire, P., & Leclere, M. (2014). Strategy repetition in young and older adults: a study in arithmetic. *Dev Psychol, 50*(2), 460-468. doi:10.1037/a0033527

Lemaire, P., Luwel, K., & Brun, F. (2017). Does the Number of Available Strategies Change How Children Perform Cognitive Tasks? Insights from Arithmetic. *Journal of Educational and Developmental Psychology, 7*(2). doi:10.5539/jedp.v7n2p43

Linnell, K. J., Caparos, S., de Fockert, J. W., & Davidoff, J. (2013). Urbanization decreases attentional engagement. *J Exp Psychol Hum Percept Perform, 39*(5), 1232-1247. doi:10.1037/a0031139

Lipton, P. A., & Eichenbaum, H. (2008). Binding and organization in the medial temporal lobe. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behavior* (pp. 337-363). New York: Oxford Universtiy Press.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: developmental differences in learning the forms of causal relationships. *Cognition, 131*(2), 284-299. doi:10.1016/j.cognition.2013.12.010

Luchins, A. S. (1942). Mechanization of problem solving: The effect of Einstellung. *Psychological Monographs, 54*(6), 1-95.

Luchins, A. S., & Luchins, E. H. (1950). New experimental attempts at preventing mechanization in problem solving. *The Journal of General Psychology, 42*, 279-297.

Luwel, K., Schillemans, V., Onghena, P., & Verschaffel, L. (2009). Does switching between strategies within the same task involve a cost? *Br J Psychol, 100*(Pt 4), 753-771. doi:10.1348/000712609X402801

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., . . . Zhao, Y. (2014). The evolution of self-control. *Proc Natl Acad Sci U S A, 111*(20), E2140-2148. doi:10.1073/pnas.1323533111

Manrique, H. M., & Call, J. (2015). Age-dependent cognitive inflexibility in great apes. *Animal Behaviour*, *102*, 1-6.

Manrique, H. M., Völter, C. J., & Call, J. (2013). Repeated innovation in great apes. *Animal Behaviour, 85*(1), 195-202. doi:10.1016/j.anbehav.2012.10.026

Marshall-Pescini, S., & Whiten, A. (2008). Chimpanzees (Pan troglodytes) and the question of cumulative culture: an experimental approach. *Anim Cogn, 11*(3), 449-456. doi:10.1007/s10071-007-0135-y

Masley, S., Roetzheim, R., & Gualtieri, T. (2009). Aerobic exercise enhances cognitive flexibility. *J Clin Psychol Med Settings, 16*(2), 186-193. doi:10.1007/s10880-009-9159-6

Masuda, T., & Nisbett, R. E. (2001). Attending Holistically Versus Analytically: Comparing the Context Sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology, 81*(5), 922-934. doi:10.1037//0022-35I4.81.5.922

Masuda, T., & Nisbett, R. E. (2006). Culture and change blindness. *Cognitive Science, 30*, 381-399.

Mathot, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav Res Methods, 44*(2), 314-324. doi:10.3758/s13428-011-0168-7

McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *Br J Psychol, 102*(1), 1-18. doi:10.1348/000712610X493115

McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cognitive Development, 22*(3), 353-364. doi:10.1016/j.cogdev.2007.01.001

McKelvie, S. J. (1984). Einstellung: Still alive and well. *The Journal of General Psychology, 112*(3), 313-315.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, memory, and cognition, 22*(6), 1423.

Menzel, E. W., Jr., & Menzel, C. R. (2007). Do Primates Plan Routes? Simple Detour Problems Reconsidered. In D. A. Washburn (Ed.), *Primate Perspectives on Behavior and Cognition* (pp. 175-206). Washington (DC): American Psychological Association.

Michimata, C., Okubo, M., & Mugishima, Y. (1999). Effects of Background Color on the Global and Local Processing of Hierarchically Organized Stimuli. *Journal of Cognitive Neuroscience, 11*(1), 1-8. doi:10.1162/089892999563201

Miles, C., Morgan, M. J., Milne, A. B., & Morris, E. D. M. (1996). Developmental and individual differences in visual memory span. *Current Psychology: Developmental, Learning, Personality, and Social, 15*(1), 53-67.

Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper and Row.

Miller, E. K., & Buschman, T. J. (2008). Rules through recursion how interactions between the frontal cortex and basal ganglie may build abstract, complex rules from concrete, simple ones. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behavior* (pp. 419-455). New York: Oxford University Press.

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Curr Dir Psychol Sci, 21*(1), 8-14. doi:10.1177/0963721411429458

Miyamoto, Y., Nisbett, R. E., & Masuda, T. (2006). Culture and the physical environment. Holistic versus analytic perceptual affordances. *Psychol Sci, 17*(2), 113-119. doi:10.1111/j.1467-9280.2006.01673.x

Moore, R. (2016). Meaning and ostension in great ape gestural communication. *Anim Cogn, 19*(1), 223-231. doi:10.1007/s10071-015-0905-x

Moriguchi, Y., Tanaka, M., & Itakura, S. (2011). Executive function in young children and chimpanzees (Pan troglodytes): evidence from a nonverbal Dimensional Change Card Sort task. *The Journal of Genetic Psychology, 172*(3), 252-265.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353-383.

Nielsen, M., Mushin, I., Tomaselli, K., & Whiten, A. (2014). Where culture takes hold: "overimitation" and its flexible deployment in Western, Aboriginal, and Bushmen children. *Child Dev, 85*(6), 2169-2184. doi:10.1111/cdev.12265

Nielsen, M., & Tomaselli, K. (2010). Overimitation in Kalahari Bushman children and the origins of human cultural cognition. *Psychol Sci, 21*(5), 729-736. doi:10.1177/0956797610368808

Nisbett, R. E., Choi, I., Peng, K., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291-310.

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: holistic versus analytic perception. *Trends Cogn Sci, 9*(10), 467-473. doi:10.1016/j.tics.2005.08.004

Nissen, H. W., Blum, J. S., & Blum, R. A. (1948). Analysis of matching behavior in chimpanzee. *Journal of comparative and physiological psychology, 41*(1), 62-74.

Norenzayan, A., & Nisbett, R. E. (2000). Culture and Causal Cognition. *Current Directions in Psychological Science, 9*(4), 132-135.

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science, 26*, 653-684.

Oden, D. L., Thompson, R. K., & Premack, D. (1988). Spontaneous transfer of matching by infant chimpanzees (Pan troglodytes). *Journal of experimental Psychology: Animal Behavior Processes, 14*(2), 140.

Ohshiba, N. (1997). Memorization of serial items by japanese monkeys, a chimpanzee, and humans. *Japanese Psychological Research, 39*(3), 236-252.

Oishi, S., Jaswal, V. K., Lillard, A. S., Mizokawa, A., Hitokoto, H., & Tsutsui, Y. (2014). Cultural variations in global versus local processing: A developmental perspective. *Developmental Psychology, 50*(12), 2654-2665. doi:10.1037/a0038272

10.1037/a0038272.supp (Supplemental)

Perelman, P., Johnson, W. E., Roos, C., Seuanez, H. N., Horvath, J. E., Moreira, M. A., . . . Pecon-Slattery, J. (2011). A molecular phylogeny of living primates. *PLoS Genet, 7*(3), e1001342. doi:10.1371/journal.pgen.1001342

Peterson, E. R., & Deary, I. J. (2006). Examining wholistic–analytic style using preferences in early information processing. *Personality and Individual Differences, 41*(1), 3-14. doi:10.1016/j.paid.2005.12.010

Petrides, M. (2008). Selection between Competing Responses Based on Conditional Rules. In S. A. Bunge & J. D. Wallis (Eds.), *The Neuroscience of Rule-Guided Behavior* (pp. 3-22). New York: Oxford University Press.

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

Pope, S. M., Fagot, J., Meguerditchian, A., Washburn, D. A., & Hopkins, W. (Submitted). Enhanced cognitive flexibility in the semi-nomadic Himba. *Cognition*.

Pope, S. M., Meguerditchian, A., Fagot, J., & Hopkins, W. (In Prep). The evolution of cognitive flexibility in primates: Chimpanzees' intermediate susceptibility to cognitive set.

Pope, S. M., Meguerditchian, A., Hopkins, W. D., & Fagot, J. (2015). Baboons (Papio papio), but not humans, break cognitive set in a visuomotor task. *Anim Cogn, 18*(6), 1339-1346. doi:10.1007/s10071-015-0904-y

Potts, R. (2012). Environmental and Behavioral Evidence Pertaining to the Evolution of EarlyHomo. *Current Anthropology, 53*(S6), S299-S317. doi:10.1086/667704

Price, E. E., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2009). A potent effect of observational learning on chimpanzee tool construction. *Proc Biol Sci, 276*(1671), 3377-3383. doi:10.1098/rspb.2009.0640

Qiu, J., Li, H., Jou, J., Liu, J., Luo, Y., Feng, T., . . . Zhang, Q. (2010). Neural correlates of the "Aha" experiences: evidence from an fMRI study of insight problem solving. *Cortex, 46*(3), 397-403. doi:10.1016/j.cortex.2009.06.006

Reader, S. M. (2003). Innovation and social learning: individual variation and brain evolution. *Animal Biology, 53*(2), 147-158.

Reader, S. M. (2016). Animal social learning: associations and adaptations. *F1000Res, 5*. doi:10.12688/f1000research.7922.1

Reader, S. M., Morand-Ferron, J., & Flynn, E. (2016). Animal and human innovation: novel problems and novel solutions. *Philos Trans R Soc Lond B Biol Sci, 371*(1690). doi:10.1098/rstb.2015.0182

Rilling, J. K., & Insel, T. R. (1999). The primate neocortex in comparative perspective using magnetic resonance imaging. *Journal of Human Evolution, 37*(2), 191-223.

Ringo, J. L., Doty, R. W., Demeter, S., & Simard, P. Y. (1994). Time is of the essence: a conjecture that hemispheric specialization arises from interhemispheric conduction delay. *Cerebral Cortex, 4*(4), 331-343.

Ritchie, S. J., & Tucker-Drob, E. M. (2017). How much does education improve intelligence? A meta-analysis. Retrieved from https://psyarxiv.com/kymhp

Roberson, D., Davidoff, J., Davies, I., & Shapiro, L. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50*(4), 378-411.

Roberson, D., Davidoff, J., & Shapiro, L. (2002). Squaring the circle: The cultural relativity of ‘good shape’. *Journal of Cognition and Culture, 2*(1), 29-51.

Roberts, A. C. (2008). Dopaminergic and Serotinergic modulation of two distinct forms of flexible cognitive control: Attentional set-shifting and Reversal learning. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behavior* (pp. 283-312). New York: Oxford University Press.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207.

Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology, 20*, 1209-1224.

Rosati, A. G. (2017). The Evolution of Primate Executive Function: From Response Control to Strategic Decision-Making *Evolution of Nervous Systems* (pp. 423-437).

Rosenthal, R., & Rosnow, R. L. (2009). *Artifact in behavioral research*. New York: Oxford University Press.

Rumbaugh, D. M. (1971). Evidence of qualitative differences in learning processes among primates. *Journal of comparative and physiological psychology, 76*(2), 250-255.

Ruscio, A. M., & Amabile, T. M. (1999). Effects of instructional style on problem-solving creativity. *Creativity Research Journal, 12*(4), 251-266.

Rushworth, M. F. S., Croxson, P. L., Buckley, M. J., & Walton, M. E. (2008). Ventrolateral and Medial Frontal Contributions to Decision-Maing and Action Selection. In S. A. Bunge & J. D. Wallis (Eds.), *The Neuroscience of Rule-Guided Behavior* (pp. 129-157). New York: Oxford Univeristy Press.

Rygula, R., Walker, S. C., Clarke, H. F., Robbins, T. W., & Roberts, A. C. (2010). Differential contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal learning. *J Neurosci, 30*(43), 14552-14559. doi:10.1523/JNEUROSCI.2631-10.2010

Sakai, K. (2008). Task set and prefrontal cortex. *Annu Rev Neurosci, 31*, 219-245. doi:10.1146/annurev.neuro.31.060407.125642

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nat Neurosci, 6*(1), 75-81. doi:10.1038/nn987

Schenker, N. M., Hopkins, W. D., Spocter, M. A., Garrison, A. R., Stimpson, C. D., Erwin, J. M., . . . Sherwood, C. C. (2010). Broca's area homologue in chimpanzees (Pan troglodytes): probabilistic mapping, asymmetry, and comparison to humans. *Cereb Cortex, 20*(3), 730-742. doi:10.1093/cercor/bhp138

Schoenemann, P. T., Sheehan, M. J., & Glotzer, L. D. (2005). Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nat Neurosci, 8*(2), 242-252. doi:10.1038/nn1394

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. *Transfer of learning from a modern multidisciplincary perspective*, 1-51.

Semendeferi, K., Lu, A., Schenker, N., & Damasio, H. (2002). Humans and great apes share a large frontal cortex. *Nat Neurosci, 5*(3), 272-276. doi:10.1038/nn814

Shan, I. K., Chen, Y. S., Lee, Y. C., & Su, T. P. (2008). Adult normative data of the Wisconsin Card Sorting Task in Taiwan. *Journal of the Chinese Medical Association, 71*(10), 517-522.

Smaers, J. B., Gomez-Robles, A., Parks, A. N., & Sherwood, C. C. (2017). Exceptional Evolutionary Expansion of Prefrontal Cortex in Great Apes and Humans. *Curr Biol, 27*(5), 714-720.

Smith, K. S., & Graybiel, A. M. (2016). Habit Formation. *Dialogues in clinical neuroscience, 18*(1), 33.

Star, J. R., & Seifert, C. (2006). The development of flexibility in equation solving. *Contemporary Educational Psychology, 31*(3), 280-300. doi:10.1016/j.cedpsych.2005.08.001

Stevens, N. J., Seiffert, E. R., O'Connor, P. M., Roberts, E. M., Schmidtz, M. D., Krause, C., . . . Temu, J. (2013). Palaeontological evidence for an Oligocene divergence between Old World monkeys and apes. *Nature, 497*, 611-614.

Stoet, G., & Snyder, J. (2007). Extensive Practice does not Eliminate Human Switch Costs. *Cognitive, Affective, and Behavioral Neuroscience, 7*(3), 192-197.

Stoet, G., & Snyder, J. (2008). Task-Switching in Human and Nonhuman Primates: Understanding rule encoding and control from behavior to single neurons. In S. A. Bunge & J. D. Wallis (Eds.), *The Neuroscience of Rule-Guided Behavior* (pp. 227-253). New York, New York: Oxford University Press.

Stoet, G., & Snyder, L. H. (2003). Executive control and task-switching in monkeys. *Neuropsychologia, 41*(10), 1357-1364. doi:10.1016/s0028-3932(03)00048-4

Stoet, G., & Snyder, L. H. (2004). Single neurons in posterior parietal cortex of monkeys encode cognitive set. *Neuron, 42*(6), 1003-1012. doi:10.1016/j.neuron.2004.06.003

Sullivan, E. V., Mathalon, D. H., Zipursky, R. B., Kersteen-Tucker, Z., Knight, R. T., & Pfefferbaum, A. (1993). Factors of the Wisconsin Card Sorting Test as measures of frontal-lobe function in schizophrenia and in chronic alcoholosim. *Psychiatry Research, 46*(2), 175-199.

Sweller, J. (1980). Hypothesis salience, task difficulty, and sequential effects on problem solving. *The American Journal of Psychology, 93*(1), 135-145.

Sweller, J. (1983). Control mechanisms in problem solving. *Memory & Cognition, 11*(1), 32-40.

Sweller, J., & Gee, W. (1978). Einstellung, the sequence effect, and hypothesis theory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(5), 513-526.

Sweller, J., & Levine, M. (1982). Effects of goal specificty on means-ends analysis and learning. *Journal of Experimental Psychology: Learning, memory, and cognition, 8*(5), 463.

Sweller, J., Mawer, R. F., & Howe, W. (1982). Consequences of history-cued and means-end strategies in problem solving. *The American Journal of Psychology, 95*(3), 455-483.

Tebbich, S., Griffin, A. S., Peschl, M. F., & Sterelny, K. (2016). From mechanisms to function: an integrated framework of animal innovation. *Philos Trans R Soc Lond B Biol Sci, 371*(1690). doi:10.1098/rstb.2015.0195

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Philos Trans R Soc Lond B Biol Sci, 364*(1528), 2405-2415. doi:10.1098/rstb.2009.0052

Thomason, M. E., Race, E., Burrows, B., Whitfield-Gabrieli, S., Glover, G. H., & Gabrieli, J. D. (2009). Development of spatial and verbal working memory capacity in the human brain. *J Cogn Neurosci, 21*(2), 316-332. doi:10.1162/jocn.2008.21028

Thompson, R. K., Oden, D. L., & Boysen, S. T. (1997). Language-Naive Chimpanzees (Pan troglodytes) Judge Relations Between Relations in a Conceptual Matching-to-Sample Task.

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Curr Dir Psychol Sci, 18*(5), 259-263. doi:10.1111/j.1467-8721.2009.01648.x

Titz, C., & Karbach, J. (2014). Working memory and executive functions: effects of training on academic achievement. *Psychological Research, 78*(6), 852-868.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16*, 495-552.

Toni, I., Rowe, J., Klass, S. E., & Passingham, R. E. (2002). Changes of Cortico-striatal Effective Connectivity during Visuomotor Learning. *Cerebral Cortex, 12*, 1040-1047.

Trick, L., Kempton, M. J., Williams, S. C., & Duka, T. (2014). Impaired fear recognition and attentional set-shifting is associated with brain structural changes in alcoholic patients. *Addict Biol, 19*(6), 1041-1054. doi:10.1111/adb.12175

Van Leeuwen, E. J., Cronin, K. A., Schutte, S., Call, J., & Haun, D. B. (2013). Chimpanzees (Pan troglodytes) flexibly adjust their behaviour in order to maximize payoffs, not to conform to majorities. *PLoS One, 8*(11), e80945. doi:10.1371/journal.pone.0080945

van Schaik, C. P., Burkart, J., Damerius, L., Forss, S. I., Koops, K., van Noordwijk, M. A., & Schuppli, C. (2016). The reluctant innovator: orangutans and the phylogeny of creativity. *Philos Trans R Soc Lond B Biol Sci, 371*(1690). doi:10.1098/rstb.2015.0183

Visalberghi, E., & Fragaszy, D. (2002). “Do Monkeys Ape?” - Ten Years After. In K. Daughternhahn & C. L. Nehavniv (Eds.), *Imitation in Animals and Artifacts.* (pp. 471-499). Cambridge (MA): MIT Press.

Vlamings, P. H., Hare, B., & Call, J. (2010). Reaching around barriers: the performance of the great apes and 3-5-year-old children. *Anim Cogn, 13*(2), 273-285. doi:10.1007/s10071-009-0265-5

Wertheimer, M. (1945). *Productive Thinking*. New York: Harper.

Yamamoto, S., Humle, T., & Tanaka, M. (2013). Basis for cumulative cultural evolution in chimpanzees: social learning of a more efficient tool-use technique. *PLoS One, 8*(1), e55768. doi:10.1371/journal.pone.0055768

Yeung, M. K., Han, Y. M., Sze, S. L., & Chan, A. S. (2015). Abnormal Frontal Theta Oscillations Underlie the Cognitive Flexibility Deficits in Children With High-Functioning Autism Spectrum Disorders. *Neuropsychology, 30*(3), 281-295.

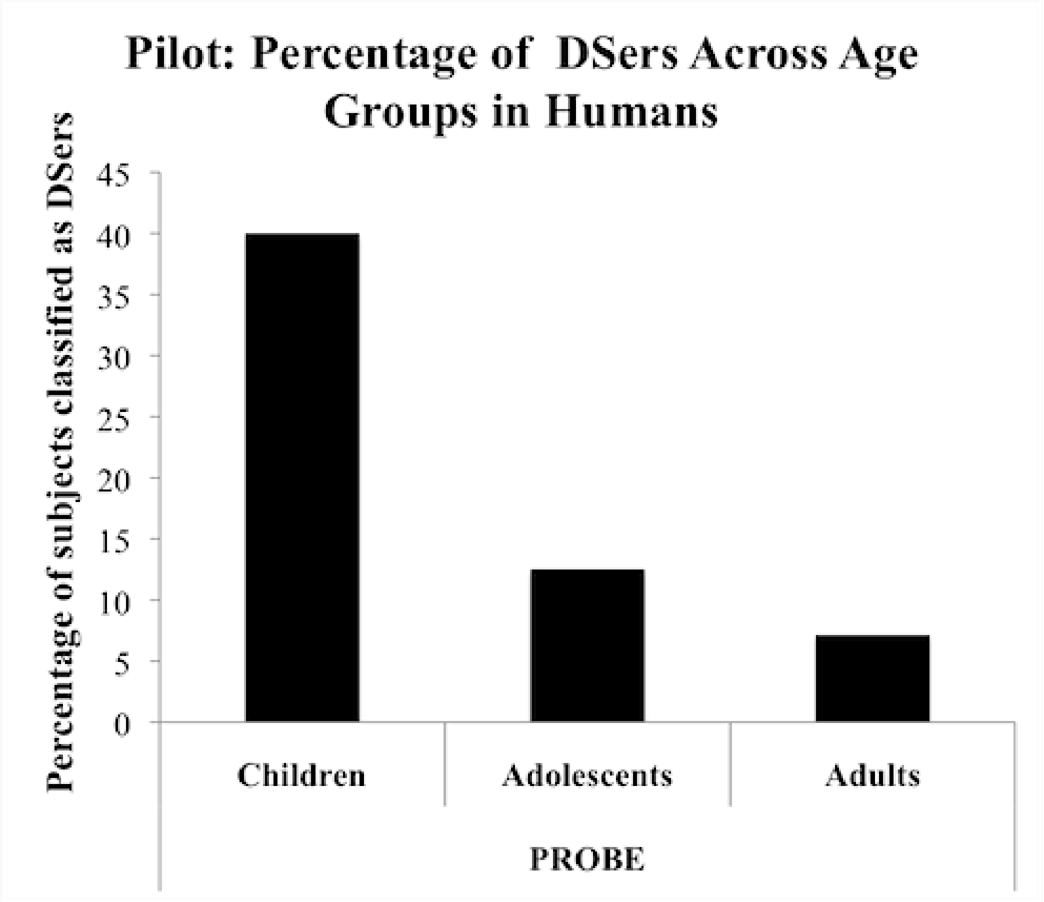
Zelazo, P. D. (2008). The Development of Rule Use in Children. In S. A. Bunge & J. D. Wallis (Eds.), *Neuroscience of Rule-Guided Behavior* (pp. 441-455). New York: Oxford University Press.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*(1), 37-63. doi:10.1016/s0885-2014(96)90027-1

Zelazo, P. D., Muller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., . . . Carlson, S. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.

APPENDICES

Appendix A



Supplemental Fig 2.1. Pilot data was collected on 32 humans (ages 6-51), including 5 children (mean age = 6.4, SD=.55), 8 adolescents (mean age = 14.13, SD=.35), and 14 adults (mean age 36.36, SD= 10.02). Methods were highly similar to those previously described; however, children were given 500ms demonstration slides during testing. Additionally, the first 10 adults were only given 48 testing trials. After a subject noted that she "figured it out at the very end," the trial numbers were doubled. Once participants had completed all trials, they were asked if they had thought about touching the Triangle directly and their responses were recorded. Our results showed that 1 out of 14 (7.14%) adults, 1 out of 8 (12.5%) adolescents and 2 out of 5 (40%) children would be classified as DSers. This is consistent with our later findings.

Appendix B - Chapter 5 Supplementary Data

Group differences in accuracy and response times

Western subjects' overall accuracy in experimental trials was significantly higher than Urban, and Traditional Himba subjects,' as revealed by a Kruskal Wallis one-way analysis of variance (H(2) = 54.74, $p < .001$; *M*s: Western = 91.5%, Urban = 77.9%, Traditional = 79.2%). Additionally, a loglinear analysis revealed no effect of sex on DS-use in any group for either PRE [χ^2(1, *N*= 183 = 0.87, $p = .351$; Female DSers = 23.6%; Male DSers = 35.6%] or POST [χ^2(1, *N* = 183 = 1.49, $p = .223$; Female DSers = 35.5%; Male DSers = 41.1%] trials.

We assessed general differences in response times between groups by isolating the time between the end of the Square1→Square2 demonstration and subjects' first response (RT1), between first and second response (RT2) and between second and third response (RT3) for BASE trials in which they used the LS. Data were normalized using a natural logarithmic transform and outliers were excluded *(N* = 6 Westerners, 1 Urban, 8 Traditional). A mixed design ANOVA indicated a significant interaction ($F(4,330) = 19.78$, $p < .000$) between response (RT1, RT2, RT3) and group (Western, Urban, Traditional). RT1 was slower than RT3, which was slower than RT2, for all groups. Further, Westerners' were significantly faster than Urban Himba, who in turn were significantly faster than traditional Himba for RT1 (Westerners: *M* = 722.6 ms; *SD* = 176.9; Urban: *M* = 1259.6 ms; *SD* = 438.6; Traditional: *M* = 1463.9 ms; *SD* = 482.3) and RT2 (Westerners: *M* = 218.1 ms; *SD* = 107.2; Urban: *M* = 405.7 ms; *SD* = 266.1; Traditional: *M* = 520.9 ms; *SD* = 225.5). For RT3, Westerners were significantly faster than Urban and Traditional Himba, who did not differ from each other (Westerners: *M* = 653.9 ms; *SD* = 115.5; Urban: *M* = 929.1 ms; *SD* = 193.9; Traditional: *M* = 980.5 ms; *SD* = 320.4).

An intermediate strategy

During analysis, we discovered that some subjects used an intermediate strategy, in which they selected Square1 and then the Triangle, skipping Square2 (Figure 1.6-1c). That is to say, they began to use the LS but then switched to the DS, invoking a ‘switch strategy’ (SS). Using the same greater than 5% classification criteria to identify SSers, we found a significantly higher proportion of SSers within the Traditional Himba (20.0%) than the Urban Himba (3.7%) however, the proportion of Westerners (5.6%) did not differ from either group, in PRE trials [$\chi^2(1, N = 183) = 10.837, p = .004$]. Yet, in POST trials SSer proportions of Traditional (16%), Urban (7.4%), and Westerners (9.3%) were not statistically distinct.