Spring 2019

A Cognitive Task as Enrichment for Chimpanzees (*Pan troglodytes*) in Sanctuary

Federico Rubio
Central Washington University, federico.rubio@cwu.edu

Recommended Citation

Rubio, Federico, "A Cognitive Task as Enrichment for Chimpanzees (*Pan troglodytes*) in Sanctuary" (2019). *All Master's Theses*. 1253.
https://digitalcommons.cwu.edu/etd/1253

A COGNITIVE TASK VIA INTERACTIVE TECHNOLOGY AS ENRICHMENT FOR CHIMPANZEES (*PAN TROGLODYTES*) IN SANCTUARY

A Thesis
Presented to
The Graduate Faculty
Central Washington University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Primate Behavior

---

by
Federico Salud Rubio
May 2019

CENTRAL WASHINGTON UNIVERSITY

Graduate Studies

We hereby approve the thesis of

Federico Salud Rubio

Candidate for the degree of Master of Science

APPROVED FOR THE GRADUATE FACULTY

Dr. Jessica Mayhew, Committee Chair

Dr. Kara Gabriel

Dr. Matthew Altman

Dean of Graduate Studies

ABSTRACT

A COGNITIVE TASK VIA INTERACTIVE TECHNOLOGY AS ENRICHMENT FOR CHIMPANZEES (*PAN TROGLODYTES*) IN SANCTUARY

by

Federico Salud Rubio

May 2019

Chimpanzees (*Pan troglodytes*) hold an especially powerful attraction for researchers interested in cognition and how it developed to the degree observed in humans specifically and primates more generally. Chimpanzees are behaviorally complex primates with compelling data supporting their possession of intricate internal lives. The objective of this study was to simultaneously learn more of the cognitive process of inductive reasoning while also assessing the efficacy of a computerized box as a novel form of enrichment to aid in improving the environment of captive chimpanzees. Three social groups at the retired medical research chimpanzee sanctuary Project Chimps were given voluntary access to an interactive box inspired by previous touchscreen studies. Individuals varied widely in their interactions with the box but did not perform above chance on the preliminary levels of the task. Interest in the box was also observed to correlate negatively with the number of sessions so that exploration in the final sessions were significantly lower than that recorded in the first sessions. These results highlight the need for refinement in provisioning captive chimpanzees with interactive technology as enrichment.

ACKNOWLEDGEMENTS

Sufficient gratitude cannot be expressed toward all those who contributed to make this project a reality. Dr. Mayhew’s support of this unorthodox project was a refreshing allowance to creatively investigate the frustrating-to-quantify phenomenon that is known as cognition. Her knowledge and expertise opened the can of worms that is cognitive research and helped create the scaffold upon which this project was built. Thanks are also due to my other committee members Dr. Kara Gabriel and Dr. Matthew Altman, whose perspectives outside anthropology were invaluable in establishing a solid foundation for this work. A special thanks to Mike Budd, without whose knowledge of electronics and shop equipment this project would not exist. Thank you to Project Chimps, not only for hosting me and this project, but for allowing me to witness the growth of what will surely be a gold standard of chimp welfare. My appreciation also goes out to Harriet Gallu, who tirelessly assisted me in gathering my initial behavioral data. Thanks also go out to Scott Zerkle, who lent a hand when the electronics tried to get the better of me. Thank you to Anthony Denice for being a champion ambassador between CWU’s primate program and Project Chimps. A very big thank you to my programmer Mike Hooper, whose knowledge of electronics finally brought these boxes to life. Finally, thank you to the Prentice Fund and Graduate Studies of CWU for the award and Summer Research Fellowship Grant. Without such assistance this project could not have come into fruition.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

Table Page

LIST OF FIGURES

APPENDIX A – SANCTUARY AND BOX PHOTOS

APPENDIX B – SUPPLEMENTAL TABLES AND FIGURE

Figure Page

CHAPTER I

INTRODUCTION

The discussion of cognition and its degree in nonhuman animals only deepens with each new examination into the workings of the mind. Historically, animals were believed to be no more than biological automatons whose behavior could be attributable to the programming of their primitive instincts (see Allen & Trestman, 2017, for a review). The discoveries of the mechanisms of evolution and genetics contribute to the understanding that all life on the planet is intrinsically connected to all other life and established the context by which we can interpret the behavior and cognition of nonhuman animals. Through this context, we have developed the view that life exists on a continuum, and that the features we use to define ourselves as human are in fact shared by other species, most especially by our closest living relatives, the primates. This understanding also shapes our view on the morality of captivity and animal welfare, and what obligations we have to the beings kept in our care.

The tribe Hominini was once a diverse clade but over time was reduced to our own single human species, *Homo sapiens.* This loss of diversity left a distinct evolutionary gap between humans and other primates, which has proven to be a great frustration for those seeking answers to the when, why, and how humanity came to be so distinct among animals. Because of this lack of extant hominin species, chimpanzees (*Pan troglodytes*) have become a focal species in inter-species comparative investigations – a direct result of their genetic proximity to our own species. Humans and chimpanzees are famously proclaimed to share approximately 99% of the same DNA (Cohen, 2007). This genetic similarity directly resulted in chimpanzees becoming a favored model in medical and pharmaceutical research in the United States, a practice that did not officially end until recently (Grimm, 2015). The exact degree of relatedness between ourselves and our next of kin has been scrutinized on many occasions, and some estimates of relatedness have suggested that our two species are not quite so genetically similar as has been previously purported (Cohen, 2007). Bonobos (*Pan paniscus*) share approximately the same degree of relatedness to humans as chimpanzees do, having diverged from chimpanzees between two and a half million or possibly eight hundred thousand years ago (Stone et al., 2010). However, bonobos are far less numerous than chimpanzees and, as such, were not the preferred model for comparative research.

The debate over whether the genus *Pan* should be included in the tribe Hominini has become a sensationalized issue that will not soon be resolved (Goodman et al., 1998). In terms of comparative cognitive studies with animals in general, whether it is worthwhile to draw comparisons between our cognition and that of another species yields several viewpoints. Comparing the cognition of any two species that occupy different ecological niches and face incomparable selective pressures raises concern over the validity of those comparisons, and interspecies comparisons become even more muddled when considering how captive environments may affect the ontogeny of cognition in contrast to wild-born/reared counterparts (Boesch, 2007).

The defining features of humans that have traditionally been used as criteria to elevate humanity above others of the animal kingdom (i.e., tool use, reasoning ability, culture, language, episodic memory) are constantly redefined, expanded, or in some instances completely eliminated as knowledge is gathered about the nature and abilities of nonhuman animals. Culture (Whiten et al., 1999; van Leeuwen, Cronin, & Haun, 2014), episodic memory (Taylor, 2014), language (Radick, 2007), reasoning ability (Premack & Premack, 1994), and tool use (Limongelli, Boysen, Visalberghi, & Snowdon, 1995; Bond, Kamil, & Balda, 2003; Pruetz & Bertolani, 2007; Vaesen, 2012; Taylor, 2014) have all been showcased in species other than our own. As the line that distinguishes humanity from nonhuman animals blurs, one concept becomes clear: nonhuman primates are complex creatures with complex needs, and it is our duty as their custodians to fulfill those needs, especially in their captive environments.

Being our closest genetic relatives alongside the bonobo, chimpanzees have been made the subject of numerous cognitive studies examining a wide range of cognitive functions including reasoning, working memory, and even theory of mind. A primary objective for many of these studies is to delineate which cognitive abilities we share with our closest living relatives and if so, to what degree. Once this objective is met, the goal then becomes to apply these delineations to the evolutionary framework to gain a better understanding not of what makes our species unique but how it became unique.

CHAPTER II

LITERATURE REVIEW

Facets of Intelligence

Cognition is the summation of mental processes that constitute thought, memory and learning. Intelligence can be considered the adeptness with which these processes occur and is derived from modular, or domain-specific, cognitive abilities (Byrne, 2000; review: Shettleworth, 2001). These modular, domain-specific cognitive abilities can be attributed to specific locations in the brain known as cortical areas, and specialization of these areas is what distinguishes different types of cognitive processes (e.g., working memory vs. muscle or procedural memory, visual vs. olfactory processing, etc.) (Bear, Connors, & Paradiso, 2007). Because these are discrete areas of the brain and because individual species are subjected to unique demands on their survival, there is great diversity in the evolution and complexity of these areas among species and therefore a great diversity in intelligence (Bear et al., 2007).

Complex behaviors are occasionally interpreted erroneously as being indicators of higher intelligence, such as in beehive construction. When these behaviors are inflexible and are not versatilely applied to a variety of situations, they are more likely attributable to being specialized adaptations of evolution and not individual intelligence (Tomasello & Call, 1997). Another potential misinterpretation of behavior occurs with associative learning, a relatively simple cognitive mechanism by which animals form correlations among objects or stimuli in their environments. This type of learning does not represent a fundamental understanding of the relationship between the objects and stimuli, only the outcomes of those relationships.

Memory is the process by which past experiences can be recalled and is a cognitive ability that is not restricted to one cortical area of the brain (Bear et al., 2007). Episodic memory “receives and stores information about temporally dated episodes or events and temporal-spatial relations among these events” (Tulving, Donaldson, & Bower, 1972, p. 385). In other words, it is the memory that allows one to recall events from the past in contrast to muscle, or procedural, memory, which is subconscious and related to proprioception. Episodic memory has also been referred to as mental time travel or being able to imagine oneself in scenarios occurring in either the past or the future.

Curiosity, or neophilia, which is most concisely defined as the seeking of new information, is a personality trait that is ancillary to intelligence and can be advantageous in discovering novel exploitable resources (Damerius, Graber, Willems, & van Schaik, 2017). This is the basis for innovation, which is preserved across time via social learning between conspecifics through horizontal and vertical transmission (Ramsey, Bastian & van Schaik, 2007; van Schaik, & Burkart, 2011). Curiosity was likely the catalyst for the development of spear hunting of bushbabies in the chimpanzees of Fongoli (Pruetz & Bertolani, 2007) and termite-fishing and ant-dipping in Gombe (McGrew, 1979), and as such, may be a key prerequisite to hypotheses on the origin of intelligence such as the social and cultural intelligence hypotheses (van Schaik, & Burkart, 2011). It has been proposed that the underlying motivations of curiosity are to either reduce feelings of uncertainty about new objects or to stimulate arousal because it is pleasurable (Litman, 2005). This latter hypothesis is the one of interest in terms of providing novel enrichment to captive animals because it insinuates that novelty in and of itself is stimulating and, therefore, beneficial for captive welfare (Forss, Schuppli, Haiden, Zweifel, & van Schaik, 2015). Neophobia is the reverse of neophilia and is hypothesized to have developed as a strategy to avoid possible threats, such as by discouraging ingestion of unknown foods that could be lethal. Animals in captivity are not immune to neophobia, and this aversion should be kept in mind in considerations of welfare as well (Gustafsson, Saint Jalme, Bomsel, & Krief, 2014).

Transitive inference is the formation of a relationship between two entities based on their relationships to a common factor and is presumed to have developed in social species with linear dominance hierarchies. This type of reasoning allows individuals to infer linear relationships between two of their conspecifics based on the individuals' relationships with a third individual (Vasconcelos, 2008). Investigations into this particular cognitive mechanism and evidence for its pervasiveness encompass a diverse array of species including honey bees, pigeons, rats, platyrrhine and catarrhine monkeys, and chimpanzees (Vasconcelos, 2008). Transitive inference ability requires some degree of abstract thinking in that an individual that demonstrates this ability is capable of imagining relationships based on incomplete information. This capacity to infer relationships without explicit information was essential to the successful completion of the cognitive task presented in this study.

Metacognition is the cognitive ability "to mentally access, evaluate, and act upon one's body of knowledge," or in lay terms, the process of thinking about thinking (Neldner, Collier-Baker, & Nielson, 2015, p. 683). This cognitive ability is one of the more challenging abilities to quantify and measure. Designing experiments to test for metacognition is complex and requires careful execution to ensure that positive results are in fact evidence of metacognition and not a result of other more parsimonious explanations (Marsh & MacDonald, 2012). In the context of primatology, questions of metacognition arise in social interactions among individuals. For chimpanzees specifically, Hare, Call, Agnetta and Tomasello (2000) demonstrated that chimpanzees are cognizant of what their subordinate or dominant conspecifics perceive visually and select food items based on that perceived perception. A subsequent experiment by the same investigators modified the original experiment by providing one individual with more knowledge of the whereabouts of the food than the other (Hare, Call, & Tomasello, 2001). The results of this study strongly suggest that chimpanzees are aware of the knowledge their conspecifics possess and make strategic decisions utilizing this knowledge of another's knowledge. Further yet, another study provided evidence that chimpanzees are capable of knowing that others make inferences about hidden food rewards and modify their decisions based on those inferences (Schmelz, Call, & Tomasello, 2011).

Theory of mind describes the ability to attribute mental states, such as motives and intentions, to oneself and other individuals (Premack & Woodruff, 1978). Experiments with mouse controllers and computer screens have shown that chimpanzees are aware of their own self-agency when moving the cursor (Kaneko & Tomonaga, 2011) but a more difficult question to answer is whether they are able to ascribe agency and minds to others. Gaze-following is often used as an indirect indication of theory of mind. Individuals that follow the gaze of their conspecifics know that their conspecific is perceiving something of interest and then attempt to take on that perspective by following their gaze (Call, Hare, & Tomasello, 1998). This interpretation suggests that the gaze follower has attributed a mental state to another and, therefore, possesses a theory of mind (Tomasello, Call, & Hare, 2003). Hare et al. (2000) ultimately conclude in their subordinate/dominant visual knowledge study that chimpanzees are indeed capable of knowing what their conspecifics do and do not see, which in turn supports the argument for theory of mind in chimpanzees but thus far has not been demonstrated to equal that of humans (Tomasello et al., 2003). These studies were followed by another that asserted that although chimpanzees are aware of what others know, they do not possess the mental faculties to understand when others hold false beliefs (Kaminski, Call, & Tomasello, 2008). This experiment was a derivation of the false belief studies with human children, which showed that human children are not capable of understanding that another person can hold a belief that is inconsistent with reality until around 4 years old (Wellman, Cross & Watson, 2001). Other methods for assessing theory of mind include experiments testing whether primates make decisions based on their knowledge of another individual's knowledge, which has yielded some interesting results supporting theory of mind in apes (Krupenye, Kano, Hirata, Call, & Tomasello, 2016).

The facet of intelligence of interest in the current study is the capacity for inductive reasoning. Induction, or the ability to infer general conclusions from specific examples, is a form of reasoning that becomes especially pertinent in foraging. Being able to infer the location of food based on seasonality, distance, and number of conspecifics present allows individuals to effectively allocate their time to searching for and consuming food, as demonstrated by yellow-nosed monkeys (*Cercopithecus ascanius whitesidei*) (MacDonald, Wilkie, & Gallup, 1990), cotton-top tamarins (*Saguinus oedipus*) (Banerjee et al., 2009) and chimpanzees (Bonnie et al., 2012; Hopper, Kurtycz, Ross, Bonnie, & Vonk, 2015). The purpose of the current study was two-fold: to assess to what degree our closest living relative shares this ability with humans, while also providing access to novel cognitive enrichment in their captive environments.

Hypotheses on the Origin of Intelligence

A higher degree of intrinsic intelligence in humans is asserted in the general intelligence hypothesis and has historically been cited as a key difference between our species and others (Reader, Hager, & Laland, 2011). However, it is no longer argued that general intelligence is exclusive to humans after examinations across multiple species and domains of intelligence (Herrmann, Call, Hernàndez-Lloreda, Hare, & Tomasello, 2007; Banerjee et al., 2009; Reader et al., 2011; Matzel & Sauce, 2017). These studies demonstrate that, although nonhuman animals do not possess the same facets of intelligence to the same degree as humans, some of them do share, if not exceed, its core rudiments in several areas (e.g., object permanence, spatial memory, reasoning, theory of mind) that cannot be disregarded or interpreted as insignificant (Herrmann et al., 2007; Marino et al., 2007; Banerjee et al., 2009; Taylor, 2014).

The social intelligence hypothesis postulates that the complexity inherent to social interactions acts as a driving force for the evolution of greater intelligence and cognitive ability (Jolly, 1966; Cheney & Seyfarth, 1990). Social animals face intricate interactions with their conspecifics daily that vary based on the unique relationship with that specific individual, and as such are under greater pressure to develop cognitive abilities that allow them to understand and navigate these complex interactions. These cognitive abilities are observable in the neural mechanisms and physical structures of certain brain regions known to be associated with social interaction (Platt, Seyfarth & Cheney, 2015). How social intelligence might increase survival and fitness led to the creation of cooperative task paradigms, which became a widely popular test of social intelligence in numerous species such as Bornean orangutans (Chalmeau, Lardeux, Brandibas, & Gallo, 1997), tufted capuchins (Chalmeau, Visalberghi, & Gallo, 1997), tufted capuchins and chimpanzees (Visalberghi, 1997), bonobos (Hare, Melis, Woods, Hastings, & Wrangham, 2007), elephants (*Elephas maximus*) (Plotnik, Suphachoksahakun, & de Waal, 2011), chimpanzees (Schneider, Melis, & Tomasello, 2012) and finally wolves and dogs (*Canis lupus* & *Canis lupus familiaris*) (Marshall-Pescini, Schwarz, Kostelnik, Virányi, & Range, 2017). Criticisms have arisen against the results obtained from chimpanzee involvement in these types of studies, however, given the ecological invalidity of forcing chimpanzees to cooperate when the social structure of the species is more conducive to competition (Hare et al., 2001; Schmelz & Call, 2016).

Diet, specifically frugivory, is another prospective selective pressure on the evolution of intelligence (Decasien, Williams, & Higham, 2017). High seasonality of fruit availability forces the animals that depend on those resources to be temporally and spatially aware of that availability (Byrne, 2000), thereby creating pressure for advanced temporal and spatial memory, which is consistent with findings in corvids that cache their food (Taylor, 2014). Food availability and its potential effect on intelligence emphasizes ecological rather than social roots for evolved cognitive abilities. No one hypothesis is likely to be the sole selective pressure for intelligence in any species when considering the diverse ecological niches that species occupy (Dunbar & Shultz, 2007). This becomes most apparent when observing the high degree of cognitive ability in animals such as cephalopod molluscs, which subsist in an ecology and possess a nervous system that are completely foreign to that of our own (Ikeda, 2009).

Animal Cognition

Investigations into animal cognition span across countless taxa, and clades that are known for their above average cognitive ability tend to share the common factor of a higher brain-volume-to-body ratio such as in cephalopods (Ikeda, 2009), corvids (Bond et al., 2003; Taylor, 2014), cetaceans (Marino et al., 2007), and several primates (Byrne, 2000; Reader & Laland, 2002; Reader et al., 2011). Which volume/ratio measures of the brain best reflect accurate predictions of higher cognitive ability and whether the ratio itself is an appropriate indicator of these abilities are debatable topics (Byrne, 2000), but some corvids do tend toward a high brain volume to body ratio and are known for higher instances of tool use than any other bird group (Bond et al., 2003; Taylor, 2014). Ravens (*Corvus corax*) have demonstrated the ability to recollect which experimenters will give a more preferred food reward a month after initial trials (Müller, Massen, Bugnyar, & Osvath, 2017). Additionally, scrub jays (*Aphelocoma coerulescens*) that cache food have been shown to understand the degradation rates of different food items and visit cache sites based on these expiration dates to maximize their usage of stored food, suggesting an ability to form and recollect episodic-like memory (Taylor, 2014). In addition to this capacity for recollection, jays are also capable of anticipating future food availability and cache current food according to the degradation rates of the soon-available resources, thus providing evidence for their possession of forethought (Taylor, 2014).

Further evidence for the exceptional cognitive abilities of corvids involved the capture of American crows (*Corvus brachyrhynchos*) by experimenters wearing masks

(Taylor, 2014). When the same crows re-encountered individuals wearing the same masks of their capturers, they reacted aggressively, suggesting that they retained the memory of the transgressions performed by the masked experimenters. A more parsimonious explanation states that the crows learned through association that the masks were linked to the threat of capture, and they reacted accordingly. Because we do not yet possess the means to quantitatively measure thought and are incapable of providing evidence for cognitive convergence, it is vital to avoid equating behavioral convergence with cognitive convergence. Positive results should be interpreted cautiously (Taylor, 2014).

Despite these shortcomings, the signature-testing approach can be used to identify explicit indicators of cognitive mechanisms via their errors, biases, and limits in an attempt to provide empirical evidence for cognitive convergence between corvids and other non-related taxa (Taylor, 2014). Invasive methodologies and their moral consequences aside, behavioral examinations of cognitive mechanisms provide the best means to contrast cognitive convergence. Standardization of these kinds of experiments would be invaluable in comparative studies of cognition and will facilitate identification of parallel cognitive mechanisms that were convergently evolved to satisfy the same needs as other species of different taxa and ecologies.

Spatial Cognition

Spatial cognition relates to how an organism navigates its environment based on its knowledge of the physical composition of that environment. This type of cognition is used on a daily basis as it is the way in which animals are able to locate or return to important feeding and resting sites (for review see Call, 2000; Garber & Dolins, 2014) as well as the means by which they predict the future position of objects in motion (Washburn & Rumbaugh, 1992). Being able to travel to feeding sites efficiently is essential because inefficient travel that yields little or no nutrition puts more caloric strain on the traveling animal (review: Call, 2000; Hopper et al., 2015). In one experiment of reversal learning and spatial memory, yellow-nosed monkeys searched for food among eight cups distributed around their enclosure, only half of which were baited (MacDonald et al., 1990). Once the monkeys retrieved all the food rewards, the other half of cups that had not originally contained food were baited. The monkeys were able to find the correctly baited cups quickly and accurately in four or five selections, even after an hour had passed between trials (MacDonald et al., 1990). Primates in the wild tend to take straight paths along familiar travel routes to established high-yield feeding sites and some spider monkeys (*Ateles belzebuth*) have been shown to choose more efficient routes than computer simulation models would predict (Garber & Dolins, 2014).

Temporal Cognition

Sumatran orangutans (*Pongo abelii*) and chimpanzees have been assessed for the cognitive ability of forethought (Osvath & Osvath, 2008). This experiment allowed for the testing of the Bischof-Köhler hypothesis on mental time travel, which asserts that nonhuman animals lack this ability. In this study, individuals were trained to retrieve a liquid food reward with a straw-like tool. During the experimental trials, individuals were given the option to select the straw-like tool from three pseudo-randomly placed distractors to eliminate biases caused by placement of the tools. They were then allowed to keep the tool they had selected when returning to their main enclosure area. After 70 minutes, they were brought to the testing area where the reward was kept. Subjects could access the reward if they both chose and retained the appropriate tool. Correct selection was 100% in all but one trial for one of the chimpanzees. Each individual was only given fourteen trials, as a way to minimize development of associative learning, but associative learning does not require excessive repetition and can be established with as little as one instance (Bernstein & Webster, 1980).

Constant rewarding via the same tool meant that associative learning could be the likely explanation for success in this study. This process of a reward reinforcing a behavior is known as operant conditioning, and is the same method used in positive reinforcement training to promote desired behaviors (Pomerantz & Terkel, 2009). In order to rule out operant conditioning and associative learning as potential explanations for the results observed in this experiment, the paradigm could have been modified to tease apart what features of the straw-like tool the subjects chose that tool for. Removing the first straw-like tool and replacing it with one that accomplished the same goal would provide support for the orangutans and chimpanzees comprehending the nature of the problem and its correct solution. Despite uncertainty in the cognitive ability of the individual subjects, “the lack of observable hesitation in the tool use was striking and would suggest a high fidelity in the envisioning of the function and the necessary manipulations” (Osvath & Osvath, 2008, p. 671).

Tool-Use

Tool-use tasks are a commonly used method to gauge the mental capabilities of nonhuman animals. Wolfgang Köhler’s 1925 experiments with chimpanzees were among the first explorations into the cognitive ability of apes with tools (as cited by Tomasello & Call, 1997). In one set of experiments, a chimpanzee was given interlocking sticks that when connected allowed the chimpanzee to reach distant food through a fence. The chimpanzee initially attempted to reach the bananas with only half of the connecting sticks to no avail. Upon manipulation of the sticks and realizing that they could be connected, the individual immediately realized the value of the length of the newly fashioned tool and returned to the fence to retrieve the banana. In a subsequent set of experiments by Köhler in 1926, bananas were suspended out of immediate reach of the chimpanzees, and in order to reach them, individuals had to stack several boxes on top of each other and then climb the makeshift tower to acquire the food. Through these experiments, Köhler hoped to find evidence for insight learning in chimpanzees, or the ability to solve a problem without stumbling upon a solution by trial and error (Ash, Jee, & Wiley, 2012). Köhler's chimpanzees were eventually successful in both tasks, but critics argued that the interlocking of sticks and stacking of boxes by the chimpanzees may have been learned techniques from previous experiences and, therefore, did not provide evidence for insight learning.

One study of tool use involved the use of a trap tube in which chimpanzees could only retrieve a food reward if they pushed it away from the hole in which it would then be irretrievable with a stick (Limongelli et al., 1995). This same paradigm was initially used to test tufted capuchins (*Cebus apella*), but it was ultimately concluded that the capuchins did not have a comprehension of what constituted a successful solution (Visalberghi, Limongelli, & Gallup, 1994). The transparency of the tube and trap meant that the chimpanzees had the opportunity to evaluate the puzzle before attempting to retrieve the reward. An immediate understanding of the solution was referred to as "a priori representational strategy” (homologous to Köhler’s insight learning) and trial-and-error attempts as “ongoing action anticipatory strategy” (Limongelli et al., 1995).

Performance between chimpanzees varied greatly, with some showing no clear grasp of the properties of the apparatus and others demonstrating a strong comprehension of what led to successful acquisition of the food reward. One individual held their hand under the opposite side of where they pushed the food reward, demonstrating their anticipation of the food reward falling out, but did not fully grasp the nature of the apparatus and, therefore, left several trials empty-handed. Orangutans were given a similar task in an examination of curiosity and its relationship with problem-solving ability but were less successful at this task compared to the other tasks (Damerius et al., 2017). Ultimately, these studies substantiate that the primates tested did not understand the task sufficiently enough to form a mental representation of the problem and were incapable of solving the task quickly and consistently through that representation. Forming a mental representation of the task in the current study was critical to consistent successful completion as well.

Broad Assessments

The Primate Cognition Test Battery (PCTB) (Tomasello & Call, 1997) created a standardized assessment of cognitive ability across multiple species. In one execution of the battery 105 human children, 106 chimpanzees, and 32 Bornean orangutans (*Pongo pygmaeus*) were tested by Herrmann et al. (2007). The battery is divided into physical and social categories. The physical domain of the test battery scores subjects’ comprehension of spatial displacements, quantity discrimination ability, and understanding of causal relations between objects. Examples include object permanence, or recognizing continuity of existence even when an item is out of view, as well as transposition, the ability to track an object as it changes location. Tool use causality was examined as well and is defined as adeptness at using tools to achieve goals, which necessitates a comprehension of cause and effect to do so. The social domain assessment tests the ability to learn socially (i.e., imitation, or the process of observing a behavior by a conspecific and adopting that behavior), understanding of communicative cues, and searches for evidence of theory of mind by gaze following.

Human children outperformed chimpanzees in terms of object permanence, whereas chimpanzees were better than children at object transposition (Hermann et al., 2007). All three species were adept at recognizing larger quantities but chimpanzees were more skillful at combining quantities. Children performed better than chimpanzees or orangutans in three of the causality tests, but chimpanzees were better than human children and orangutans at tool use causality. Human children outperformed both ape species in the social domain overall, with social learning being their area of greatest proficiency. These results provide strong evidence for the cultural intelligence hypothesis, an extension of the social intelligence hypothesis, which claims that human hyper-sociality created even greater selective pressure for intelligence (Herrmann et al., 2007). These results do not support the general intelligence hypothesis. Were this the case, human participants should have outperformed the other two ape species across all domains. However, it can be argued that the young age (around 2.5 years old) of the subjects was an important factor in the results obtained. Many cortical areas are not yet fully developed in children (Bear et al., 2007), and therefore children are not a valid sample to compare against the performance of adult organisms. Once again, the validity of interspecies cognitive comparative studies is dubious, and the expectation of primates to operate at the same tier as human children is reminiscent of a *scala naturae* interpretation of life. Support for general intelligence as a shared ancestral primate characteristic has been demonstrated in the cognitive ability of cotton top tamarins, which is particularly interesting because they reside in a relatively distant cladistic branch in relation to apes (Banerjee et al., 2009).

Two cercopithecine monkeys, olive baboons (*Papio anubis*) and long-tailed macaques (*Macaca fascicularis*), were also given the PCTB to assess the hypothesis that intelligence and brain size are positively correlated (Schmitt, Pankau, & Fischer, 2012). Both species demonstrated skill in the physical domain, particularly in quantity discrimination. Neither species generated significant results in social learning, but they were still adept at reading communicative cues. Compared to previous studies (Herrmann et al., 2007), the long-tailed macaques outperformed both chimpanzees and orangutans and the olive baboons outperformed orangutans in gaze following. According to the authors, these results do not necessarily indicate that cercopithecine monkeys demonstrate more evidence for theory of mind than chimpanzees and orangutans; rather, subject apathy in the chimpanzee and orangutan participants may have occurred (Schmitt et al., 2012). Another issue with gaze following is whether it is indicative of theory of mind or is merely an associative behavior or orientation reflex. Additionally, differences in gaze following could be explained by differences in socio-ecology rather than phylogeny (Schmitt et al., 2012). Overall, there were no clear distinctions between the monkey and ape responses. The lack of interspecies differences suggest that the tests may not be refined enough to discern subtle differences in cognitive ability (Premack &

Woodruff, 1978; Call & Tomasello, 2008; Schmitt et al., 2012) but the alternative remains that each of these species genuinely do not differ greatly across these domains.

Nonmatch-to-Sample

In psychology, the nonmatch-to-sample (NMS) paradigm is a well-studied experimental design in which a subject must acquire a reward by picking a response that does not match the initial sample or location, and so the selection of novelty is what constitutes a correct response. Most NMS studies are conducted with delays between initial sample exposure and selection in order to test memory, such as in aging rats (Chrobak, Hanin, Lorens, Napier, & Squire, 1995). Performance in NMS can vary greatly between subjects and is considered to be cognitively demanding given that subjects must select responses that are unlike their sample cues (Elliott, Norris, Ettlinger, & Mishkin, 1977; Falcone et al., 2013). This necessitates the cognitive ability to abstractly categorize two objects as being similar or dissimilar to each other, a fundamental prerequisite for success in the task administered in the current study.

Oddity

Making distinctions and recognizing oddity or identifying objects that do not match the other objects in a set is crucial because it can mean the difference between a nutritious meal and a lethal mistake when selecting foods for consumption in wild contexts. Identifying oddity is intertwined with inductive reasoning in that it involves recognizing patterns and objects that do not fit those patterns, as well as making inferences about the objects that do not adhere to those patterns. Oddity recognition studies have been conducted in chimpanzees using human eye gaze direction (Tomonaga, Imura, & Emery, 2010). In primates, prolonged eye contact can incite agonistic behavior, or combative social interactions, in certain contexts and it is for this reason that many primates are hypothesized to have darkened sclera surrounding the iris of the eye, to obscure the individual's gaze direction and mitigate unintended aggression (Kobayashi & Kohshima, 2001). This is in contrast to the non-pigmented sclera of humans. An analysis of gaze following in apes and human infants provides support for the cooperative eye hypothesis, in which it is proposed that the white sclera of the human eye allows easier gaze following to improve joint attentional interactions and cooperation (Tomasello, Hare, Lehmann, & Call 2007). Which aspects of the face are crucial in gaze discrimination have been evaluated using both simple (e.g., visual-search methods and polarity experiments) and advanced methods (e.g., eye-tracking technology) (Tomonaga et al., 2010). In their experiments that manipulated the orientation, gaze direction and color polarity of photographs of human faces, Tomonaga et al. (2010) found indications that chimpanzees are more adept at identifying direct gazes than averted gazes. In the context of the current study, being able to identify oddity would allow individuals to succeed in six out of the seven levels of the task if they were able to form a generalized mental representation.

Analogical Reasoning

Analogical reasoning studies offer an enlightening area of nonhuman primate cognitive exploration. Another form of inductive reasoning, analogical reasoning studies offer the opportunity to gauge whether animals understand the abstract relationships between objects. Gillan, Premack, and Woodruff (1981) tested the chimpanzee Sarah for this ability in a series of experiments. Trials were separated into analogy completion, in which Sarah would choose from several options to form an appropriate analogy based on an example analogy, and analogy identification, in which she was to select the comparable pair to her example analogy from several dyads of symbols. Sarah achieved significant success in each experiment, which gave the researchers cause to believe that at least some chimpanzees are capable of understanding some forms of analogical reasoning. The encultured life history of Sarah was likely a key factor in her success, since she had spent a significant amount of time performing cognitive tasks with humans prior to this specific study (Gillan et al., 1981).

Four Guinea baboons (*Papio papio*) were given similar analogical matching-to-sample tasks through computerized touchscreens (Fagot & Maugard, 2013). The baboons were to match responses to the sample based on its shape, color, both shape and color, or neither in probe trials. The samples were generated from six colors and six shapes, and correct responses by the baboons yielded a piece of dry wheat as a reward. An 80% success stipulation over one hundred trials in two consecutive blocks was required of the baboons in order for them to progress to subsequent trials, and as a result, training averaged 58,541 trials per individual over five to six weeks. On average, individuals were 77% correct. In the next experiment, ten blocks of 128 randomized trials composed of training and probe trials were used. The baboons were 63.4% correct for probe trials. The rate of success can be interpreted to be the consequence of the extremely high quantity of training that the baboons received, although the authors maintain that the data provide evidence for the ability to match relationships across different domains.

Human, chimpanzee, and rhesus macaque (*Macaca mulatta*) subjects were tested via relational match-to-sample tasks to create a frame of reference for aptitude in analogical reasoning among different primate taxa (Flemming et al., 2008). Each subject was tasked with matching an analogical response consisting of a pair of items that matched the relationship of the items in the sample given. The human participants were tested using meaningful and non-meaningful words, the four chimpanzees were tested using known and unknown lexigrams, and the five macaques were tested with Arabic numerals. Human success surpassed both chimpanzee and rhesus macaque performances. Chimpanzees varied in their individual ability to select the correct analogies, and the macaques did not achieve success that could not be explained by chance. The chimpanzees involved in the study did not match the success observed in a similar study with the chimpanzee Sarah (Gillan et al., 1981), again possibly as a result of her uniquely encultured life history, but it was also postulated that another critical variation between the chimpanzees tested and Sarah was that Sarah had been trained with and utilized a token that she used specifically to identify objects as same or different (Flemming et al., 2008). This further illuminates the standardization issue that makes comparative data interpretation difficult, as several variables may not be held constant and have confounding effects, leaving room for misinterpretation. Similar to the other experiments referenced, this specific task exemplified how critical a mental representation of an abstract principle was for success.

Transitive Inference

Transitive inference between two corvids (pinyon jay, *Gymnorhinus cyanocephalus*; western scrub jay, *Aphelocoma californica*) was examined to search for the potential connection between sociality and intelligence (Bond et al., 2003). Both species are distinct in their sociality, with pinyon jays living in groups ranging from 50 to 500 individuals and scrub jays restricted to pair-living. The transitive inference paradigm was tailored for easy interaction for the jays and entailed the use of different colored lights coupled with pecking keys that the jays could make a selection with. Each light was associated with a certain amount of food reward, and both species learned the relationship A < B at comparable rates, but subsequent reward dyads were learned more quickly by the pinyon jays. The reward dyads in the linear hierarchy were intermixed so that the subjects were required to simultaneously track six dyadic relationships. Accuracy on the task was comparable between the two species, but pinyon jays continued to improve more rapidly than the scrub jays. Scrub jay subjects were supplemented with 100 extra training trials in an attempt to induce higher performance, but still reached an asymptotic level of performance despite these attempts (Bond et al., 2003).

The subjects were then tested on their selection in non-adjacent pairs. Both species demonstrated significantly high levels of accuracy, and no significant difference was observed between the species. Notably, each pinyon jay responded correctly with at least 70% accuracy across all pairs. Accuracy and latency differences between the two species are believed to have been the consequence of different representational strategies used by the two species (Bond et al., 2003). Both species were successful, but the more social species had a higher tendency to quickly grasp the parameters of the experiment, thus lending support to the social intelligence hypothesis.

Transitive inference studies in three juvenile chimpanzees possibly shed some light onto the ontogeny of transitive inference for their species (Gillan, 1981). In these trials, food rewards were placed in several differently colored containers, and the chimpanzees were trained until they consistently chose the larger amounts. Food rewards increased in amount across the six containers, which were labeled A through F. If chimpanzees are capable of transitive inference, we would expect them to choose the greater food reward containers in non-adjacent pairs (i.e., choosing the fourth container, D, over the second container, B). Only one of the chimpanzees chose D in the B-D pairing in 100% of trials. An explanation of this success was that the chimpanzee chose D over B not because of transitive inference ability, but rather because of the more recent reward association with D. During training, container D was not only more recently learned than B but had also yielded a larger amount.

This association interpretation was further bolstered when, in another round of experiments, the values of the colored containers were shifted so that container A became the highest value and container F the lowest. In this condition, the initially successful chimpanzee's performance suffered when selecting between non-adjacent pairs, suggesting that she had relied on an association formed early in the initial trials. However, success in the first trials implied that the subject had formed a mental linear representation of the value of each container because she was never presented with all six containers at once and had still managed to recognize the increasing value of each successive container. Learning ability was clearly demonstrated, but re-learning ability was not. A potential source of variation in task performance was the age and rearing history of the chimpanzees (Gillan, 1981). Being juveniles, the cortical areas responsible for transitive reasoning may not yet have been fully developed, which is supported by the fact that the two unsuccessful chimpanzees were the youngest of the trio. However, non-naturalistic life histories of the individuals in a captive setting in conjunction with inexperience may have affected the development of certain cognitive abilities and produced results that would not be observed in chimpanzees from wild populations.

Whether sociality, and accordingly the social intelligence hypothesis, had any bearing on transitive inference ability was tested in two lemur species: group-living ring-tailed lemurs (*Lemur catta*) and the less gregarious mongoose lemur (*Eulemur mongoz*) (MacLean, Merritt, & Brannon, 2008). Both species were tested on transitive inference through a touchscreen computer monitor to evaluate the assumption that the more social ring-tailed lemurs would be more proficient at the task as predicted by the social intelligence hypothesis. MacLean et al. (2008) found that the ring-tailed lemurs were more proficient at the task than the mongoose lemurs but, ultimately, the two lemur species did not differ fundamentally in transitive inference ability because mongoose lemurs improved their responses when cues were ordinal. Differences in spatial cognition between the species may be the result of distinct adaptive niches. However, ring-tailed and mongoose lemurs utilize their environments similarly and are therefore not likely to have significant differences in spatial cognition. Performance variability in these two species was ultimately attributed to differences in social structure rather than ecology or training performance (MacLean et al., 2008). Kittler, Schnoell, and Fichtel (2015) provided a review of cognitive studies of ring-tailed lemurs in comparison to other lemur species and concluded that, although each species performed comparably in each task, the social aspect of ring-tailed lemurs adds a level of cognitive complexity that is best represented by their communicative cues with their group mates.

Reasoning by Exclusion

Assessments for inductive reasoning led to a paradigm that would later become a standardized method. Human children and chimpanzee subjects watched two opaque cups be baited with two different foods (Premack & Premack, 1994). Following a two-minute delay, subjects observed an experimenter eating one of the two previously hidden food items. Subjects were then prompted to choose one of the cups and, if they inferred correctly based on their observations of what the experimenter had been eating, they should choose the cup that contained the food that did not match what the experimenter had been eating because it was not being eaten by the experimenter and should therefore still be hidden. This same paradigm has been modified to test species such as dogs (Erdőhegyi, Topál, Virányi, & Miklósi, 2007), Caledonian crows (*Corvus moneduloides*) (Jelbert, Taylor, Gray, & Call, 2015), apes and spider monkeys (*Ateles geoffroyi*) (Hill, Collier-Baker, Suddendorf, & Burghardt, 2011), human children (Hill, Collier-Baker, Suddendorf, & Call, 2012), as well as between goats and sheep (*Capra aegagrus hircus* and *Ovis orientalis aries*) (Nawroth, von Borell, & Langbein, 2014). Of the 20 human children tested, 18 chose the container opposite of what the trainer had chosen (Premack & Premack, 1994). The chimpanzees varied greatly in their selections. One chimpanzee chose the reward cup (still baited) in 100% of trials, whereas another chimpanzee chose the non-rewarded cup (no longer baited) in 100% of trials. Another chimpanzee did not begin to consistently select the rewarded cup until after the first two trials, and the other until after four trials.

In the final test of this experiment, the experimenter ate an item that was identical to one of the bait foods but was wrapped and therefore distinct from the food item that was hidden under one of the cups (Premack & Premack, 1994). In this condition, the older 4-year-old children picked either of the baited cups at chance levels, indicating that they comprehended that the experimenter was eating food that could not have been the same food they watched be hidden. However, 3-year-old children and the chimpanzees continued to make their selections on the assumption that a food item had been removed from one of the cups and continued to select cups that held the food item different from what the experimenter was eating.

An expansion on this original study considered age and species effects on inferential reasoning and suggested that the chimpanzees previously tested may have made their selections by association and automatically chose the opposite of what the experimenter was eating as opposed to fundamentally understanding the experiment (Call, 2006). Each nonhuman great ape species was represented in this updated study (Bornean orangutans; chimpanzees; gorillas, *Gorilla gorilla*; bonobos). Subjects varied in their success rates, and in the third experiment, subjects with poor memory retention and success were excluded. Conditions with plastic chip rewards in place of food were linked to poor performance, likely due to the absence of a food motivator. After analyzing the results across the conditions, the author concluded that subjects were not learning by association, that the ability to make inferences appeared to increase with age, and that there was no apparent evidence of species differences in inferential ability (Call, 2006).

This reasoning by exclusion paradigm (Call, 2006; Premack & Premack, 1994) has also been used to test the inferential ability of grey parrots (*Psittacus erithacu*s) in which exclusion was defined as “the ability to base a decision on the elimination of potential alternatives” (Pepperberg et al., 2013, p. 272). Process of elimination is the core of induction, and an exclusion study also creates a corollary object permanence experiment because the subjects make decisions based on their belief of whether a food reward remains after observing an experimenter with one of the rewards (Pepperberg et al., 2013). One hundred percent of the grey parrots (*n* = 4) were immediately successful in the baseline visible removal of food rewards, with 50% being completely successful in the invisible condition when the food reward was removed out of their sight, and the remaining 50% were successful in most trials.

Conditions and sides in which food rewards were placed were randomized to prevent inadvertent cueing. The errors committed by the subjects were interesting because they occurred after a series of successful selections. It is possible the parrots were engaging in exploratory behavior to discern whether the food reward had really been removed from the cup (Pepperberg et al., 2013). There are inherent obstacles in comparing the ability of distantly and even closely related species given the number of variables that cannot be controlled for such as differing ecologies or perceptive ability (Pepperberg et al., 2013). This is an especially relevant concern given the vast evolutionary distance between psittacine birds and primates but, nevertheless, cross-taxa comparisons of reasoning ability offer insight into the convergence of these skills.

The results obtained in the inference by exclusion paradigm have led researchers to investigate the specific aspects of the paradigm that lead to successful responses. Differences in perception were tested between domesticated dogs, bonobos, and chimpanzees (Bräuer et al., 2006). Dogs are known for their comprehension of human gestural pointing, in direct contrast to chimpanzees, bonobos, and wolves (Bräuer et al., 2006). The social conditions of the experiments involved the use of communicative cues, which could either be a brief point, a sustained point until a selection was made, a brief look, or a sustained look until selection of a container. A behavioral cue was another signal in which the subject watched the experimenter attempt to reach for and open the correct container. Causal cues were ones involving physical properties of the containers.

The auditory causality condition constituted the shaking of the containers by the experimenter when it was baited or empty. Another condition was the shaking of the cup via string, or even the production of an artificial sound in the cup to control for noise. The visual causality conditions replaced the baiting containers with boards that would be displaced if on top of a food reward. In one condition, the boards were baited in front of the subject, and in another they were baited before the subjects arrived. In the control for smell, both boards were displaced but only one was baited.

The dogs performed significantly better in the social cue conditions, whereas the apes were better at the causal conditions (Bäuer et al., 2006). Together these findings support the social-dog/causal-ape hypothesis, which highlights the evolutionary histories of the two species. Dogs were bred to be cooperative companions of humans and are, therefore, sensitive to the communicative cues of humans. Interestingly, human-reared wolves have demonstrated this sensitivity to human communication as well (Heberlein, Turner, Range, & Virányi, 2016). In contrast, apes, albeit social, have had less selective pressure for symbolic communication with their conspecifics and are, therefore, more dependent on other communicative cues to meet their basic needs for survival (Bäuer et al., 2006). Evolution is frugal and does not result in adaptations, physical or otherwise, that are not essential for survival in the ecological niches of the species tested, and as such, this study also emphasizes the importance of ecological context when interpreting results of failure or success in cognitive tasks.

Animal Welfare

One purpose of the current study was to improve the conditions of captivity for chimpanzees that were retired from invasive biomedical research. Although sanctuaries do their utmost to make captivity as comfortable and stimulating as possible, the animals in their care are undeniably deprived of the liberties and associated benefits experienced by living in complex wild environments. For chimpanzees that have spent extensive lengths of time in captivity, the possibility of reintroduction into wild populations is not an option. These individuals, although more than likely capable of sustaining themselves on wild forage with the proper training, would represent a threat to local human populations surrounding reintroduction sites because of their habituation to human presence and therefore higher likelihood of interaction (Hockings, Yamakoshi, Kabasawa, & Matsuzawa, 2010). To compound the infeasibility of reintroduction, widespread habitat destruction throughout Africa means that populations already residing there will face increasing density challenges as the habitats dwindle (Lee, 2010). To reintroduce more individuals would in effect compromise the quality of life for all chimpanzees living in the area (Hernández-Pacheco, et al., 2013; Allgas, et al., 2017). Additionally, biomedical research chimpanzees represent a biohazard as a result of their exposure to pathogens in laboratory settings and could jeopardize the health of wild populations and the human communities in their vicinity. They are sometimes specifically isolated when released in wild areas because of this risk (Pacelle, 2016). With all these limitations to bear, captivity is the only option for these individuals and as such it is our responsibility as their stewards to ensure that their environment is as stimulating and enriched as possible.

Improving the conditions in which animals are kept has become a top priority in captive institutions in the last several decades (Bloomsmith, Brent, & Schapiro, 1991; Broom, 1991). Historically, animal enclosures in zoological parks were little more than barren cages. Creating enclosures that reflected the habitats in which the animals are naturally found was the first step toward welfare-centric captivity (Carenzi & Verga, 2009). Enclosure size and design have shown to be influential on the behavior of the animal occupants of that area, highlighting the necessity for a higher degree of conscientiousness in housing animals when optimal welfare is the goal (Neal Webb, Hau, & Schapiro, 2018). To achieve optimal welfare, programs should be consistently monitored so that behavioral changes, whether good or bad, can be catalogued and their root causes identified. In this regard, surveys of captive primate facilities are invaluable in that they showcase what practices are commonplace and which can be improved upon (Baker, 2016). Optimal welfare can be compromised either by the presence of aversive stimuli or the lack of stimuli that would be encountered by the animal in a natural setting (Morgan & Tromborg, 2007) and can even be affected by such trivialities as the predictability of husbandry schedules throughout the course of the day (Bassett & Buchanan-Smith, 2007).

Positive reinforcement training (PRT) is a method in which associative learning is utilized to communicate to animals when they perform a desired behavior by providing a reward. The potential benefits of PRT include improved caretaker/chimpanzee relations, increased prosocial behavior, decreased abnormal behavior, decreased stress during medical procedures, reduced shifting time between enclosures, decreased agonism during feedings, and even increased reproductive behaviors (Pomerantz & Terkel, 2009). Pomerantz and Terkel (2009) were interested in the implementation of PRT and whether its use helped mitigate anxiety in 12 captive chimpanzees.

PRT reduced rates of abnormal behavior, and its use had an overall positive net effect, but its effectiveness varied between individuals (Pomerantz & Terkel, 2009). Three of the chimpanzees showed a nearly 5% decrease in the frequency of stress-related behavior (i.e. 15% pre-training, 10.2% with training). However, the dominant male showed a 55.2% increase in stress-related behavior. Prosocial behaviors generally increased across the group after PRT, with the exceptions of the alpha male and the oldest individual, potentially because of neophobia in these two specific chimpanzees (Pomerantz & Terkel, 2009). These results reiterate that universal solutions to improve welfare are a fantasy because individual variation means that uniform reactions to enrichment and training methods is unattainable. As such, enrichment must be customized not only to species but to the individual as well to achieve maximum welfare. Despite the lack of universal benefits from the use of PRT, it still offers a useful technique by which individuals can become acquainted with novel forms of enrichment with which they are otherwise reluctant to interact.

Enrichment is the addition of objects and activities to a captive animal's restricted environment to promote its expression of natural behaviors (Broom, 1991). Enrichment can be classified into five non-mutually exclusive categories of social, physical, nutritional, occupational, and sensory (Bloomsmith et al., 1991). Social enrichment includes the interaction of individuals with other individuals that are typically of the same species. Physical enrichment can be achieved in the construction of an animal's enclosure, such as designing ample room and substrate for climbing and brachiation when housing primarily arboreal primates. Nutritional enrichment involves the use of food and is typically presented in the form of extractive foraging which usually entails the animal using some degree of skill to acquire the food (e.g., simulated termite mound fishing in chimpanzees). Occupational enrichment is any enrichment that engages the animal in constant activity and is also well-represented by extractive foraging. Successful sensory enrichment can be created by providing perceptually stimulating items in the enclosure, such as mirrors or video recordings (Bloomsmith & Lambeth, 2000).

The ultimate goal of enrichment is to provide adequate stimulation in the restricted environment of captivity, thereby mitigating any adverse psychological or behavioral consequences that can develop as a result from the deprivation of that stimulation. Abnormal behaviors (e.g., stereotypical behavior, self-injurious behavior, regurgitation and re-ingestion, etc.) are well-documented in captive animals and are distinct from normal behaviors by their frequency and prevalence in captive animals in contrast to wild populations of the same species (Birkett & Newton-Fisher, 2011; Broom, 1991; Hook et al., 2002; Lopresti-Goodman, Bezner, & Ritter, 2015; Walsh, Bramblett, & Alford, 1982). Greater frequencies of abnormal behavior in captive versus wild populations highlights the need for environmental enrichment that arouses captive animals into performing more species-typical behavior and, thereby, increases their well-being (Broom, 1991).

What behaviors can be accurately classified as abnormal has proven to be a contentious topic. Initial ethograms of abnormal chimpanzee behaviors included such actions as coprophagy, fecal smearing, and stereotypic movements (Walsh et al., 1982). Arguments have been made against the default consideration of stereotypical behaviors and coprophagy as absolute indicators of compromised welfare (Mason & Latham, 2004; Hopper, Freeman, & Ross, 2016) because their classification as such may be symptomatic of an anthropocentric view of what constitutes appropriate normal behavior for other species. The argument against such classification being that behaviors that are unpalatable to our own species should not by default be considered inappropriate for that reason alone.

The frequency of abnormal behaviors varies between populations of captive animals, suggesting a social transmission component in the expression of abnormal behaviors, and has been found to correlate directly with environmental differences in housing (Hook et al., 2002). Jacobson, Ross, and Bloomsmith (2016) recently conducted an evaluation of the rates of abnormal behaviors in captive chimpanzees across 26 zoos. Their results showed that female chimpanzees were 3.57 times more likely to engage in coprophagy than males, which is consistent with other rates of socially learned behaviors that female chimpanzees acquire from their mothers (Hopper et al., 2016). These results further support the claim that coprophagy can more appropriately be classified as a socially learned, culturally-transmitted behavior. The overall rate of coprophagy in chimpanzees born in laboratories was 5.33 times higher compared to those born in the wild (Jacobsen et al., 2016). This higher rate can be explained by the increased opportunity that captive chimpanzees have to participate in coprophagy compared to their wild counterparts; captive chimpanzees occupy a limited space that naturally leads to a greater density of feces than wild chimpanzees encounter. Behaviors that result in self-inflicted physical harm, such as self-biting, self-hitting, and hair plucking, are unambiguous abnormal behaviors that signal poor mental and physical welfare (Lopresti-Goodman et al., 2015). Optimal welfare is achieved not only through the decrease of self-injurious behaviors but also through the promotion of species-typical behaviors (Bennett, et al., 2014; Bennett, Perkins, Tenpas, Reinebach, & Pierre, 2016). For chimpanzees, this can entail the provision of enrichment that encourages foraging over wide areas and presenting food puzzles that require tools and extraction skills to acquire the rewards.

A survey of laboratory facilities was conducted inquiring into the behavioral management and environmental enhancement implementation of each facility (Baker, Weed, Crockett, & Bloomsmith, 2007). According to the survey data, only 20% of facilities changed their enrichment programs based on internal review decisions, whereas another 65% of facilities reported changes as a result of external site visits (Baker et al., 2007). Awareness of the effectiveness of enrichment is crucial. Providing enrichment that does not engage the animals to which it is given does not improve welfare and, consequently, leads to higher costs through wasted resources and labor (Bennett et al., 2014; Bennett et al., 2016). An assessment tool for primate environmental enrichment was created by Dutton, Pierre, Bailoo, Warkins, Michel, and Bennett (2018) in which they assigned weighted scores to commonly utilized enrichment devices to identify the most effective types of enrichment based on duration of engagement, agency in engagement, and potential for interaction. According to the adjusted scores, computer-based devices, puzzles, foraging substrates and foraging objects were ranked as the most effective enrichment (Dutton et al., 2018).

Novel enrichment and its benefits for the behavior of bonobos specifically was recorded in a study that compared the baseline behavior of eight bonobos and their current enrichment items with their behaviors during the implementation of five new types of enrichment (Csatádi, Leus, & Pereboom, 2008). These bonobos were known to experience higher social stress during the winter when confined to an indoor area and had a history of rough handling of infants that had resulted in an infant's death. Each of the five new types of enrichment were food-baited and led to increased group activity as well as decreased infant separation from its mother (Csatádi et al., 2008). This study provided support for the hypothesis that enrichment should be novel and varied so that the animals do not become habituated to the enrichment and lose interest, effectively negating the purpose of the enrichment.

The advantages of different modes of enrichment, their effects on activity budgets, and the possibility of curbing the rate of habituation were evaluated in ten Javan gibbons (*Hylobates moloch*) (Gronqvist, Kingston-Jones, May, & Lehmann, 2013). Novel objects, olfactory enrichment, and food-based foraging equipment were left for five days each in the gibbons' enclosure. The gibbons were most interested in the foraging box and the novel object boomer balls but were least interested in the scented mat enrichments. Singing rates, a prosocial behavior, were significantly increased when the gibbons had access to the foraging box and boomer balls. No significant interest was observed with the use of the scented mats or boomer balls however, demonstrating that novelty alone is insufficient to be effective enrichment and, as such, it is necessary to tailor enrichment to species-specific interests and needs (Gronqvist et al., 2013).

Cognitive research tasks and their potential corollary benefits as enrichment has been subject to skepticism, but "the link between cognitive research and captive care is not only complimentary but essential to the advancement of both fields" (Ross, 2010, p. 309). Clark (2011) provides a thorough review of past cognitive studies and their influence on the enhancement of welfare of captive great apes. In a subsequent study with a cognitive puzzle challenge, Clark and Smith (2013) found evidence of behavioral benefits for adult chimpanzees in the form of increased play behavior. Successful enrichment replicates or replaces the daily challenges of the wild counterparts of captive animals. Because wild counterparts are subjected to challenges that are solvable by naturally selected cognitive skills, not providing such challenges to captive animals can be considered tantamount to deprivation of adequate stimulation. Proper balance of challenge and solvability is a precarious task and a challenge in and of itself, in that "tasks that are unchallenging cause boredom or apathy, while those that are challenging cause anxiety if the individual does not have the skills required to complete the task" (Meehan & Mench, 2007, p. 251). When providing access to these tasks, maintaining agency of the study animal is paramount to preserve welfare because subjecting individuals to tasks that they are both apathetic toward and incapable of succeeding in can lead to frustration and consequently reduce welfare (Broom, 1991).

As mentioned previously, these tasks at times can prove to be sufficiently difficult as to elicit frustration and stress. Self-directed behaviors (SDB) are considered to be indicative of both arousal and stress in multiple primate species (Baker & Aureli, 1997; Leavens, Aureli, Hopkins, & Hyatt, 2001; Wagner, Hopper, & Ross, 2016). In their study of SDB in chimpanzees and gorillas, Wagner et al. (2016) had their subjects participate in a two-part cognitive task separated by difficulty into easy and hard. Both species performed better in the easy task than the hard task and, as predicted, SDB rates increased with the frequency of incorrect responses. Similar results were observed by Clark and Smith (2013) in their cognitive challenge device. Stress response and SDBs may not have a simple linear relationship, however, as suggested by a study with a female orangutan that exhibited lower salivary cortisol after having participated in cognitive tasks when compared to baseline cortisol levels when she did not engage in these tasks (Elder & Menzel, 2001). The lower cortisol levels provided evidence that the traditional indicators of stress in nonhuman primates may not necessarily be indicative of a physiological stress response in every instance of SDB expression.

Individual personality traits have a direct relationship with cognitive task participation and cannot be ignored in considerations of welfare (Herrelko, Vick, & Buchanan-Smith, 2012; Altschul, Wallace, Sonnweber, Tomonaga, & Weiss, 2017). One study with 11 chimpanzees demonstrated that the openness personality measure accounted for 37.2% of the variance in cognitive task participation, and that a neuroticism personality measure predicted higher rates of SDB (49.5%) (Herrelko et al., 2012). Other stress-related behaviors that can occur and that are not self-directed can include aggressive actions directed toward the enrichment itself (Koolhaas et al., 1999). In terms of the box created for the current study, these aggressive actions can include excessive repetitive pushing of buttons or attempts to forcefully move the box via pushing or pulling. Together with SDB, these behaviors can be construed as being stress-induced and are therefore appropriately identified as stress-related behaviors (SRB).

Intraspecies individual variation in performance on cognitive tasks is imperative when interpreting cognitive ability (Limongelli et al., 1995; Hopper et al., 2014). Both these studies found that individual chimpanzees varied widely in their adeptness at completing the tasks successfully. Some chimpanzees in the trap-tube task selected the correct position from which to remove the food from the trap immediately, while others either required more trial-and-error attempts or simply chose the wrong strategy and lost the opportunity to acquire the reward altogether (Limongelli et al., 1995). Hopper et al.

(2014) delved into the implications of individual variation and its influence on problem-solving success in their assessment of 36 chimpanzees. Those researchers were interested in the effects of personality, age, sex, and estrous state on rates of success. A factor analysis of 41 personality traits revealed six dimensions of chimpanzee personality: methodicalness, extroversion, agreeableness, openness, reactivity/undependability, and dominance (Hopper et al., 2014).

Each subject performed their task out of view of their conspecifics to eliminate the possibility of social learning. Two puzzle tasks with different methods of reward retrieval were used in these experiments. There was intra-individual consistency between puzzles but no effect of age or sex (Hopper et al., 2014). No significant correlations between specific aspects of personality were directly connected to success, but significant positive correlations were found between levels of openness, dominance, and methodicalness with duration of interaction with the puzzles in the male chimpanzees. Reactivity/undependability positively correlated with higher latency to engage in the task in females (Hopper et al., 2014). The researchers mentioned that, for these novel enrichment/cognitive tasks, individual personalities were less apparent than what would be observed in a social context.

Touchscreens

Touchscreens have become so pervasive in our culture in recent years that they have even diffused into the captive environments in which we house animals that are capable of interacting with them. High profile institutions that utilize such technology include the Lincoln Park Zoo in Chicago, Zoo Atlanta, and the Primate Research Institute of Kyoto University (Egelkamp & Ross, 2019; Matsuzawa, 2003; Perdue, Clay, Gaalema, Maple, & Stoinski, 2012). This technology offers an intuitive interface for the subjects to interact with a computer through virtual tasks that offer an array of activities (Fagot & Maugard, 2013; MacLean et al., 2008; Perdue et al., 2012). Besides direct cognitive assessments, these types of devices can offer sensory enrichment when programs such as painting are installed, which has been shown to lower SRB rates (Grunauer & Walguarnery, 2018).

Implementation of interactive technology as cognitive enrichment has faced skepticism because it is inherently non-naturalistic and, as such, is not ideal for enriching a captive environment (Kim-Mccormack, Smith, & Behie, 2016). However, as has been noted previously, habituation to enrichment is common if not inevitable (Gronqvist et al., 2013). Hesitation over touchscreen implementation as enrichment may also arise from methodologies used in some laboratories that isolate individuals and restrain them to compel participation (Calapai, Berger, Niessing, Heisig, Brockhausen, Treue, & Gail, 2017). The novelty and benefits from engaging with enrichment is typically short-lived and, therefore, the need for new stimuli is constant and unrelenting. Computerized interactive technologies offer a unique solution to this issue of novelty, because activities such as self-expression, visual/audio stimulation, or even social interaction through monitors operated by zoo guests are limited only by the imagination of the designer (Clay, Perdue, Gaalema, Dolins, & Bloomsmith, 2011; Kim-Mccormack et al., 2016; Mallavarapu, Bloomsmith, Kuhar, & Maple, 2013). Even after prolonged access to computerized enrichment, rhesus monkeys have shown to maintain high levels of engagement with these types of devices (Bennett et al., 2016). Despite the unnatural nature of computerized technology as enrichment, Washburn (2015) argues for its continued implementation because “Nonhuman primates are flexible. Faced with new challenges, they learn, adapt, and thrive” (p. 226).

One method by which we can infer whether cognitive experiments improve the welfare of captive animals is by comparing activity budgets between wild and captive populations. By establishing a wild population’s activity budget as the baseline for comparison, we create an ecologically valid frame of reference from which we can estimate what proportion of behaviors in a captive chimpanzee population is considered deviant from the norm and indirectly gauge how healthy it is (Yamanashi & Hayashi, 2011). The activity budget of the wild chimpanzees in Yamanashi and Hayashi’s (2011) study closely matched previous activity budgets for the same population by other researchers, and the authors noted that chimpanzees in the experimental condition had a similar resting and feeding activity budget to the wild population. These two activity budgets differed from the activity budget of the captive control group and provide evidence that cognitive experiments are effective food-based enrichment if voluntary in nature (Yamanashi & Hayashi, 2011).

Study Objectives

The purpose of the current study was to test the inductive reasoning ability in chimpanzees and to assess the efficacy of the task and the experimental box as a means of enrichment to improve captive welfare. Participants were observed during the task to collect information on degree of interaction, duration of interaction, and any indications of stress. Participation of the chimpanzees was analyzed to identify possible predictors between demographic characteristics of the chimpanzees and box interaction.

Objective 1.

The first objective was to examine whether chimpanzees demonstrate the ability to inductively reason via the computerized task presented in the experimental box. I predicted that the chimpanzees would demonstrate evidence for this ability and that individuals would vary in their adeptness at the task.

Objective 2.

The next objective was to evaluate whether the rate of usage of the box indicates that it is an effective form of enrichment for the chimpanzees. Devices of this kind require a relatively large initial investment, and thus it is valuable to determine if the value of the box is worth the investment. I predicted that enrichment of this kind is of value if executed properly.

Objective 3.

Another aim of the current study was to determine if demographic factors such as age, sex, social group, or placement of the experimental box affected the frequency of box interaction by the chimpanzees. I predicted that not all individuals would interact with the box equally and some demographic factors would influence box interaction more than others.

Objective 4.

The final objective was to determine if potential frustration originating from using the box outweighs the potential benefits. I predicted that any expressions of stress-related behaviors associated with frustration from box use would be mild and, therefore, not preclude the box as a form of effective enrichment.

CHAPTER III

METHODS

Ethics Statement

This study was performed with three captive chimpanzee social groups. It was a non-invasive observational study in which chimpanzee participation was voluntary. Approval and permission for this study was given by both the Central Washington University Institutional Animal Care and Use Committee (protocol #A031801) and the Project Chimps Science Committee.

Study Site

Project Chimps is a nonprofit sanctuary for retired medical research laboratory chimpanzees and is funded in large part by the Humane Society of the United States. The sanctuary has access to 236 acres of forested land in the mountains of Morganton, Georgia. Founded in 2016, it is the newest chimpanzee sanctuary in the U.S. The facilities in place were originally constructed by the former Gorilla Haven, a private wildlife management operation that housed single male gorillas of Zoo Atlanta before the gorilla species survival plan was amended to house males socially. Gorilla Haven formally transferred ownership to Project Chimps in 2015. Chimpanzees began to arrive at the sanctuary in 2016 from the University of Louisiana at Lafayette's New Iberia Research Center for medical research.

The space that the chimpanzees occupy is compartmentalized to allow the separate social groups shared access to the six-acre outdoor Peachtree Habitat enclosure (hereafter referred to as the habitat) (Appendix A1). The habitat is bisected by an electrical fence corridor which permits two separate social groups to occupy the habitat at any given time. Habitat access is rotated between groups each day. Four villas are integrated into the 4.5m-tall concrete wall surrounding the habitat. Cantilevered electrical wire tops the wall and electrical fencing is used to make a secondary barrier around the villas to create a patio in which caregivers can work while chimpanzees are within the habitat (Appendix A2). Each villa is composed of two porches and two or three bedrooms. The porches are large areas outfitted with ample climbing substrate and wire mesh outer walls to allow exposure to outdoor weather conditions, while the bedrooms are fully indoor spaces to provide climate-controlled conditions. Two bedrooms are adjoined by a squeeze cage which is used for chimpanzee transfers and medical procedures. Chimpanzees gain access to the habitat from the porches through a wire mesh tunnel that runs through the patio. All doors between chimpanzee-occupied areas are remotely and electrically operated.

Subjects

The three social groups included in the study comprised a total of 26 individuals ranging in age from 7 to 27 years old (Appendix B1). Age groups of the chimpanzees were dichotomized for simplified statistical analysis via generalized linear mixed effects models. Males 14 years and younger were categorized as non-adults, whereas males 15 years and older were classified as adults; females 12 years and younger were classified as non-adults, whereas females 13 and older were adults. This age-sex classification follows Nishida et al. (2003) for the chimpanzees of Mahale. The nine chimpanzees in the Dorothy Jo & Tilly Villa (DJT) were subadult and juvenile males. Six adult males lived in the Chimps Ahoy Villa (CA), and nine adult and subadult females lived in Villa Four (V4). Villas house at maximum two social groups.

It is not unusual for chimpanzees in medical research laboratories to come from diverse backgrounds. Some individuals may be wild-caught, some may be relinquished pets, and some may have only ever lived in laboratory conditions. Explicit life histories of each individual are not consistently or formally conveyed to Project Chimps but breeding records do contain information regarding parentage of younger individuals bred in captivity. Many of the chimpanzees at NIRC and Project Chimps have a filial relationship with several individuals as a result of captive breeding. The chimpanzees arrive and live in same-sex social groups until introductions can be made with an opposite-sex group. At the beginning of the study, the sanctuary had 39 chimpanzees in six social groups in residence; at the end of the study, six months later, there were 59 individuals in seven groups. One male and one female group became fully integrated over the course of the study. Group introductions are a major endeavor at Project Chimps, which aspires to form the largest social groups possible. Captive breeding is not a goal for Project Chimps and as such measures are taken to prevent the likelihood of pregnancies. The age of the chimpanzees currently in residence ranges from 7 to 35 years old.

Boxes

I designed and built by hand two computerized boxes with consultation from Imperial Woodworking in Thorp, WA. The faces of the boxes had an approximate surface area of 0.37 ßm^2 and were 20 cm deep (Appendix A3a). Only one box was given to a social group at a time. In the event that competition for box use led to agonistic behaviors between individuals as evidenced by aggressive behaviors such as threats or attacks, both boxes would be presented for use. This never became necessary. The boxes consisted of nine LED-illuminated interactive buttons corresponding to each combination of the colors red, yellow, and blue with the shapes circle, triangle, and square. These buttons were manufactured for classic arcade video game boxes and were purchased from Arcade World UK (www.arcadeworlduk.com) (Heysham, UK). Button-pressing was the means by which the chimpanzees responded during trials. The interactive buttons were randomly arranged in a circle equidistant from each other with an LCD screen in the center which displayed the color and shape cues presented during each trial. A tenth white circle LED button was built into the bottom left corner of the box and acted as the start trial button. The start button was included to measure response latency during the first trial of a session and to offer each participant the same starting point. Transparent plexiglass covered all non-interactive surfaces to reduce potential physical damage to the box as well as to seal the electronic components from liquid damage. Bamboo shoots were fastened along the front edges of the box to create a gap between the surface of the box and the mesh upon which it would hang. This was intended to create space for the chimpanzees to reach buttons in the event that the buttons did not align precisely with the gaps in the mesh.

The computer running the program was a Raspberry pi 3 microcontroller. This and other Raspberry pi accessories were purchased from SparkFun electronics (www.sparkfun.com) (Niwot, USA). The LCD screen was extracted from a pi-topCEED desktop display and was plugged directly into the Raspberry pi; the display also plugged into a second printed circuit board (HUB) that regulated power and other functions (Appendix A4). The HUB was extracted from the pi-topCEED as well. A third circuit board called a pi-top PROTO was also connected directly to the Raspberry pi 3 and was the interface between the pi and buttons. Buttons were wired directly to the pi-top PROTO. The fourth circuit board held an MCP23017 GPIO expander chip, which was required to control illumination of the LEDs in the interactive buttons. The fifth printed circuit board was a stepper motor HAT, which connected to the pi-top PROTO and interfaced with the STEPPERONLINE stepper motor that turned the gumball mechanism responsible for dispensing food rewards.

The gumball mechanism was a 10 cm circle cut from 1.2 cm plywood with eight 1.2 cm holes symmetrically cut into the circle which were large enough for only one food reward pellet. The gumball mechanism was physically connected to the stepper motor which turned 45° with each correct response so that just one food pellet hole passed over the slot that led to the food pellet tray where chimpanzees could retrieve the reward. This allowed for one pellet to be dispensed at a time. A 10-cm poly vinyl chloride plastic tube cap was used as a container for the food pellets and sat directly above the gumball mechanism. This allowed the food pellets to be gravity fed into the holes of the gumball mechanism whenever a pellet had just been dispensed. The food tray was cut from a PVC tube and was reinforced with metal brackets so that it could not be broken off by the chimpanzees. Four eyebolts were placed on four corners of the box so that the box could hang on the mesh from the top and be secured to the mesh by a ratchet strap woven through the eyes of the bolts across the bottom.

A Raspberry pi V2 camera module was placed just above the LCD screen and took a photo with each response during trials to assist in identifying which chimpanzees were engaging with the program. A programmable speaker was placed below the display, was wired to the pi-top PROTO and set to emit a chime with correct responses. The

Raspberry pi, stepper motor HAT, and HUB each required their own direct DC power, thus a power strip was installed in the box. The power strip plugged into an extension cord that entered the box through a 6-cm hole in the bottom of the box. To prevent the chimpanzees from being able to grab the extension cord, a removable PVC tube was fixed to the underside of the box that angled the cord away from the mesh and eliminated the possibility of a chimpanzee being able to pull the cord toward the mesh.

Trial Programming

A Raspberry pi runs a Linux OS and utilizes the programming language Python. I drafted the specifications for the "chimp.py" program created to execute the trials and these specifications were translated into code by a programming freelancer hired through the Freelancer website (www.freelancer.com). Rounds of trials were defined as an engagement with the box by a chimpanzee from the pressing of the start button to their leaving of the box; these engagements could occur multiple times over the hour. Trials were defined as the individual prompts/responses that occurred within a round. A trial began when a chimpanzee pressed the illuminated start button, which had a matching symbol on the display (Appendix A3a). Once a trial began, the cues displayed on the screen indicated which response buttons were incorrect and would, therefore, not yield a food reward pellet when pressed. The chimpanzees would then need to inductively reason that the illuminated button that was not displayed on the screen was the correct answer and would yield a food pellet. The displayed cues were randomized between each trial. I supervised all sessions to ensure immediate intervention should the chimpanzees' wellbeing be compromised through agonism, frustration from use of the boxes, or electrical threats from damage to wiring in the box. These situations never occurred.

Upon a correct response by a chimpanzee, a food pellet was dispensed automatically by the gumball mechanism/stepper motor apparatus into the food tray and was accompanied by a chime emitted by the speaker. Food pellets were ordered from Bio-Serv and were a mix of banana, chocolate, berry and piña colada flavors. Chimpanzees were anticipated to favor one particular flavor over others and, therefore, a random mix of the pellets would act as variable reinforcement to maintain long-term interest in participating in the tasks (Hulac, Benson, Nesmith, & Shervey, 2016). Incorrect responses did not yield a reward and resulted in a new trial with a different set of cues of the same level of difficulty. Corrective trials occurred when the subject made two errors consecutively; this was in an effort to offset disinterest or frustration as a result of a lack of food incentive. In these trials, if the chimpanzees were at an advanced level, the number of cues on the display would be reduced to a previous level, thus increasing the likelihood of a successful response. Corrective trials could not occur for consecutive incorrect responses in the first level (Level 0).

Level 0 constituted the least cognitively demanding task, in that seven of the nine response buttons were randomly occluded from selection and only one cue was presented on the display. Response buttons were illuminated and responsive to presses and were, thus, distinct from the occluded buttons which were neither illuminated nor responsive to presses. In these trials, subjects had a 50% chance of success and had a higher probability of comprehending the parameters of success more quickly. Level 1 increased the cue and response number by one so that two cues were on display and three responses were available for selection. Additionally, the two incorrect responses in this level shared a feature, meaning that they were either the same shape or color. This was done to facilitate the understanding that the correct response would be dissimilar from the other responses in some manner.

Participants would progress to each successive level by meeting a set criterion for their current level (i.e., ≥ 80% correct responses over at least 10 consecutive trials). Trial accuracy and latency of responses were recorded for each trial. During a session, the Raspberry pi Camera Module V2 captured an image with each response to later corroborate identity of the chimpanzees to match them to their responses. Every push of a button that was available as a response was also recorded into a table stored as a .csv file that included information about the current Level, the date and time, latency to respond, and whether the response was correct. The original program was later modified to also include which prompts were on display. Each trial in the first ten hours of box access began at Level A, but after chimpanzees began to show indications of habituation toward the box, a modified program was created that started at Level 0. The response buttons in this modified program would light up intermittently at the start screen in an attempt to attract the chimpanzees' attention. Whenever a trial was initiated and one minute had passed without any activity, the trial terminated and reset to the pre-trial state with the illuminated start button coupled with the white circle cue on the display. The chimp.py program originally contained five levels, which increased to seven levels after it was modified to attempt to offset disinterest.

Level A, similar to Level 1, consisted of three possible responses with two displayed cues (i.e., 33% chance success), but the two incorrect responses were not required to be similar in shape or color. Level B increased to five possible responses with four displayed cues (i.e., 20% chance success). Level C had seven responses available with six cues displayed (14% chance success). Level D included all nine responses with eight cues displayed (i.e., 11% chance success) (Appendix A3b). Because the correct response in the first four levels was consistently the odd one out, the chimpanzees could have developed the general strategy that, regardless of the number of responses available for selection, the correct solution was never displayed on the screen. The use of this strategy, in which "specific configural 'patterns' of multiple stimuli… serve as unique cues for discriminative responses" (Taniuchi, Miyazaki, & Siddik, 2017, p. 7) was not the method of interest in this study but is a potential outcome in any cognitive assessment.

The last level, Level E (or the Einstein level) was designed to counterbalance this general strategy should it have been employed by any of the chimpanzees. Two colors and two shapes were on the screen and every response button was available for selection. The two colors were displayed in two columns occupying half the display. The two shapes were white outlines in the other half of the screen. The chimpanzees that reached this level had to make broad inferences and exclude any color or shape that matched the cues. In this way, the chimpanzees could not rely upon their previous strategy to solve the task, and would have had to relearn how to arrive at a correct response. As an example, the colors red and blue on display with the shapes triangle and square would have meant that the only possible solution was the yellow circle. Subsequent levels of difficulty would become available only if the individual achieved 80% success at the current level; therefore, chimpanzees that were less skilled at the task were not given trials vastly beyond their current capacity.

Experimental Procedure

To become familiar with individual chimpanzees and to establish a baseline of their behavior, I observed each individual in the three social groups using two-minute focal scan samples (Altmann, 1974) over a one-hour period using an ethogram adapted from Neal Webb et al. (2018) (Appendix B2). Each of the 52 behaviors fell into broader categories: Aggressive, Locomotion, Submissive, Affiliative, Manipulate Object, Abnormal, Self-Directed, Foraging/Feeding, Inactive, Sexual and Other. The Other category included three behaviors: pant-hoot, which is not exclusively restricted to any one behavioral category; out of view, in which case the subject was not observable; and other, which was any behavior not listed in the ethogram. Recorded behaviors were supplemented with notes detailing information such as grooming partners in addition to describing other behaviors. Having a behavioral baseline is important when attempting to determine whether a new enrichment is effective because it creates a context from which one can measure whether behavioral frequencies were altered by that enrichment.

Each of the 26 individuals was randomly selected and observed for six hours for a total of 156 hours of observation between the hours of 11:30 am to 4:30 pm from June 9 to July 27, 2018. Observations were collected with the help of a research assistant, but IOR could not be established. Observations made by the assistant were dropped, leading to 134 hours of observation used for final analysis. This yielded unbalanced observations among individual chimpanzees with some being observed for 3 hours and others for the full 6 hours. Because these observations were gathered for context and are only presented via descriptive statistics, the imbalance is regarded as acceptable.

Presentations of the box occurred anywhere between the hours of 10:30 am and 4:30 pm and were provided for the chimpanzees' use from September 12 through December 9, 2018. Each social group was intended to have access to the boxes for 30 hours each. However, with the arrival of two new social groups from NIRC in November, quarantine procedures came into effect for DJT and V4. The two boxes were kept in these two villas and CA no longer received access to the box. The three females at DJT (i.e., Buttercup, Charisse & Emma) received the box for one hour before they were relocated to the fourth Cedar Tree Villa for the mixed-sex group formation. CA received the box for ten hours total and the last session for V4 was canceled because of difficulty shifting the chimpanzees, resulting in 69 total hours of presentation time.

Selection of social group, position of box, location of set-up, and time of set-up were pseudo-randomized on testing days and were contingent upon caregivers and their cleaning schedule as well as their ability to shift chimpanzees out of the area for set-up. Position of the box refers to whether the box was placed at ground level or an elevated height so that the chimps could use the box from the second level of their enclosures. Location of set-up refers to whether the box was placed inside the villa in either a bedroom or squeeze cage or outside so that it could be used in the porch area. Position and location were accounted for to determine whether they had an impact on the likelihood of chimpanzees interacting with a box and were initially designed to be balanced between conditions across all groups. With the onset of colder weather in October, plastic sheeting was installed around the mesh surrounding the porches to provide protection from the cold, and as such, outdoor location placement became untenable. Data were still collected on position and location but do not represent a balanced design between conditions.

A box was hung on the 25-cm^2 wire mesh barrier separating chimpanzee-occupied areas from human areas. A box was placed and removed after chimpanzees were shifted to another area so that the participants and I could have no physical contact of any kind. The boxes themselves were suspended by quick links on the human side of the barrier and were fixed in place by a ratchet strap across the bottom of the box to fasten it securely to the barrier. Due to the large size of the box, support bars in the mesh sometimes occluded part of the front surface of the box. The chimp.py program could be manually adjusted to shift where the cues were displayed on the screen to account for this occlusion.

When the box was placed at an elevated position, I used a rope and pulley system to lift the box to the appropriate height. A Campak Xtreme I+ UHD 4K camera was centered on the box to video record box interactions to corroborate identifications and behaviors with notes taken during interactions and to measure interaction durations. Once the box and camera set-up were completed, the chimpanzees were given access to the box for one-hour intervals. Chimpanzees retained their agency to engage with the box and, as such, were never isolated or coerced into interacting with the box. Stress-related behaviors were recorded by all-occurrence sampling during box interactions (Altmann, 1974).

Data Analysis

Data collected during the preliminary behavioral observation phase consisted of the behavior of each individual, the locations at which the individual was observed during the hour, and the time of day. I calculated behavioral frequencies by behavioral category rather than each behavior for concise data visualization. All behaviors of a category were summed (X) and then divided by the total number of observations ($N = 3660$) to determine their proportion ($\hat{p}$). All other statistical analyses were conducted in R (R Core Team, 2019). To measure inductive reasoning ability, I used Chi-square analysis on trial responses of the chimpanzees comparing observed frequencies of correctness to expected frequencies based on chance. The *a priori* expected ratio was established at 50% for Level 0 and 33% for Levels 1 and A based on the number of available response buttons. Individuals did not progress to any other levels. Only twelve trials of Level 0 and three trials of Level 1 occurred. Level 1 was reached not by successful completion of Level 0, but rather as corrective trials for Level A. As a result of the small sample size of these two levels, they were omitted from analysis. The chimpanzees did not engage with the box extensively enough to assess individual proficiency and, therefore, all trials in Level A were pooled and analyzed at the group level.

I analyzed box presentation and demographics of the individual chimpanzees to determine if these factors predicted whether some individuals were more likely to interact with the box than others. Generalized linear mixed models (GLMM) were created using the *glmer* function of the *lme4* package with villa (social group) as the random effect. Response variables in the models were box exploration, box use, and duration of interaction. Box exploration is defined as attention directed toward the box when an individual was within 30 cm of the box. This included behaviors such as looking, touching, smelling, and licking of the box without pushing the start button and beginning a trial. Box exploration records interest in the box as a physical object. Box use is defined as engagement with the program through the initiation of trials and the pushing of active buttons. Box exploration and use were regarded as binary outcomes, in which no exploration or use over the hour was designated as a 0 and some degree of interaction was signified by a 1. Duration of interaction, measured in seconds, is the summation of time exploring and using the box over the hour for each individual and was not binary in outcome. The duration GLMM was set to a Poisson distribution in the family argument of the *glmer* function, while the box exploration and box use models were set to binomial distributions.

Linear mixed models were ruled inappropriate for these analyses because box exploration and box use were binary response variables and interaction duration was count data. These types of response variables are not well-represented by a simple linear relationship. Generalized linear models of all predictors and responses were conducted with social group as another fixed effect, and the correlation coefficients of the fixed effects of these models were contrasted against the corresponding coefficients. The model with correlation coefficients closer in value among the estimates was deemed the better fitting model. The GLMM produced the closest coefficients and was selected in every case. These types of models are also advantageous in handling repeated measures coupled with unbalanced sampling (Crofoot, 2013).

The model predicting box exploration was corrected for quasi-complete separation, a phenomenon in which a response variable occurs exclusively with the same predictor factor in the data, which the model interprets as perfect prediction. This phenomenon led to abnormally large estimates and standard errors in the model. For this model, a logistic regression with Firth's correction was run with the *logistf* function of the *logistf* package. The random effect of villa was retained but run as another factor.

Box exploration and box use by the chimpanzees were distinguishable from each other by the audible clicks made by button pushing and the .csv files that were only generated during box use. Interactions in which button pushing occurred but did not yield a .csv file indicated that chimpanzees pushed inactive buttons and, therefore, these interactions were not defined as box use. Fixed effect variables of the GLMM included sex, age class (nonadult or adult), sequence of box presentation (first through thirtieth presentations), position and location of the box, the program version used, whether the chimpanzees simultaneously had access to the outdoor habitat, and if quarantine was in effect. Two GLMM models were created for the box exploration and box use response variables to identify potential differences between these two levels of box interaction. Generalized linear mixed models with sequence as a fixed effect and villa as a random effect were also conducted with box exploration and box use to identify habituation effects over time.

To determine the value of the box as an enrichment tool, interaction duration and observed SRB were evaluated by one-tailed *t*-tests. If one hypothesized that the box is an ineffective form of enrichment that would not interest the chimpanzees, average use was not expected to differ significantly from zero. Likewise, the hypothesis that using the box would not affect the frequency of SRB, these behaviors were also not expected to differ significantly from zero. Duration of box use and SRB could not be negative integers, and so I used a one-tailed test with a greater than direction parameter. Camera failure occurred during two box presentations: during the first presentation for the DJT males and the third presentation for the V4 females. Duration could not be corroborated with behavioral data and, therefore, these values are missing for these two presentations.

CHAPTER IV

RESULTS

Behavioral Observations

Observations of the chimpanzees revealed that the greatest proportion of behaviors was within the Inactive category (48%, n = 1778), which consisted of two behaviors: Inactive Alert, where the subject was passive but visibly alert; and Inactive Rest, where the subject was inactive with eyes closed (see Figure 5). The next highest proportion of behaviors were in the Foraging/Feeding category, which accounted for 15% of observations. Affiliative behaviors, such as grooming, represented 8% of observations, as did behaviors defined as Other. Behaviors in the Locomotion category were 7% of

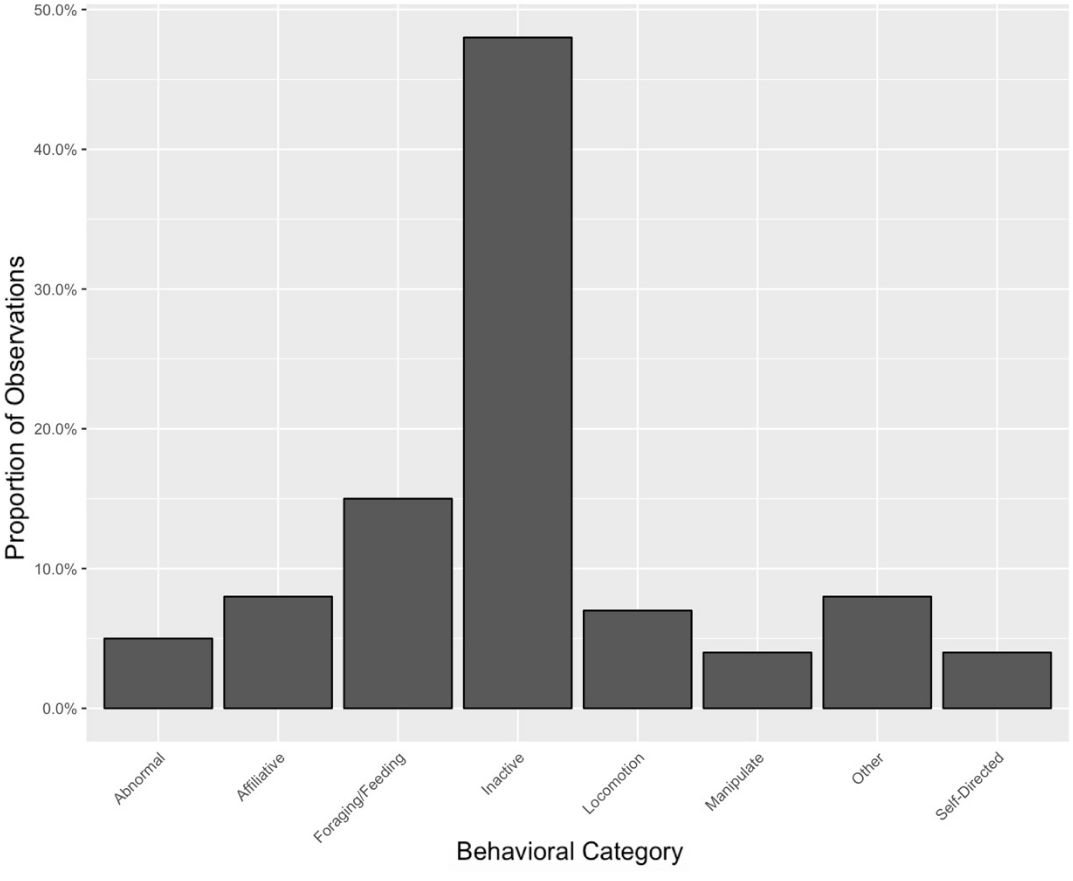


Figure 1. Proportions of observed behavior. (N = 3660)

observations, and Abnormal behaviors only represented 5% of observations. Both categories of Self-Directed and Manipulate were recorded at 4% each. Subjects were observed to engage in behaviors belonging to the Aggressive ($n = 18$), Sexual ($n = 9$), and Submissive ($n = 3$) categories, but the frequency of these behaviors did not reach the threshold of being at least 1% of total observations and, as such, were statistically represented as occurring at 0% and omitted from graphical display (see Figure 1).

Box Interactions

Over the 69 hours of box presentations, the chimpanzees explored the box 185 times. Of those total explorations, 36 led to box use by 13 chimpanzees (Appendix B3). One individual, Noel, comprised 16 of the 36 box uses (i.e., 44% of total). During those 36 instances of box use, the chimpanzees completed 186 trials. Instances in which a trial was initiated by pushing the start button but did not result in an input within one minute were classified as no response (NR) and were not considered a completed trial. Chimpanzees initiated but did not complete 25 trials across all three levels. At Level 0, chimpanzees completed seven trials and left five trials as NR. Only three trials were completed at Level 1, and at Level A, chimpanzees completed 176 trials while leaving 20 trials as NR. I performed a Chi-square analysis of Level A with chance response frequencies as 59 correct and 117 incorrect according to the 1:2 odds of guessing the correct answer. Level A responses were 55 correct and 121 incorrect ($\chi^2 = 0.34$, df = 1, $p > 0.5$) and as such were not significantly different from chance levels (see Figure 2).

Predictors of Box Interaction

The models analyzing factors predicting box exploration, box use, and duration of interaction produced results of mixed significance. The Firth's corrected model

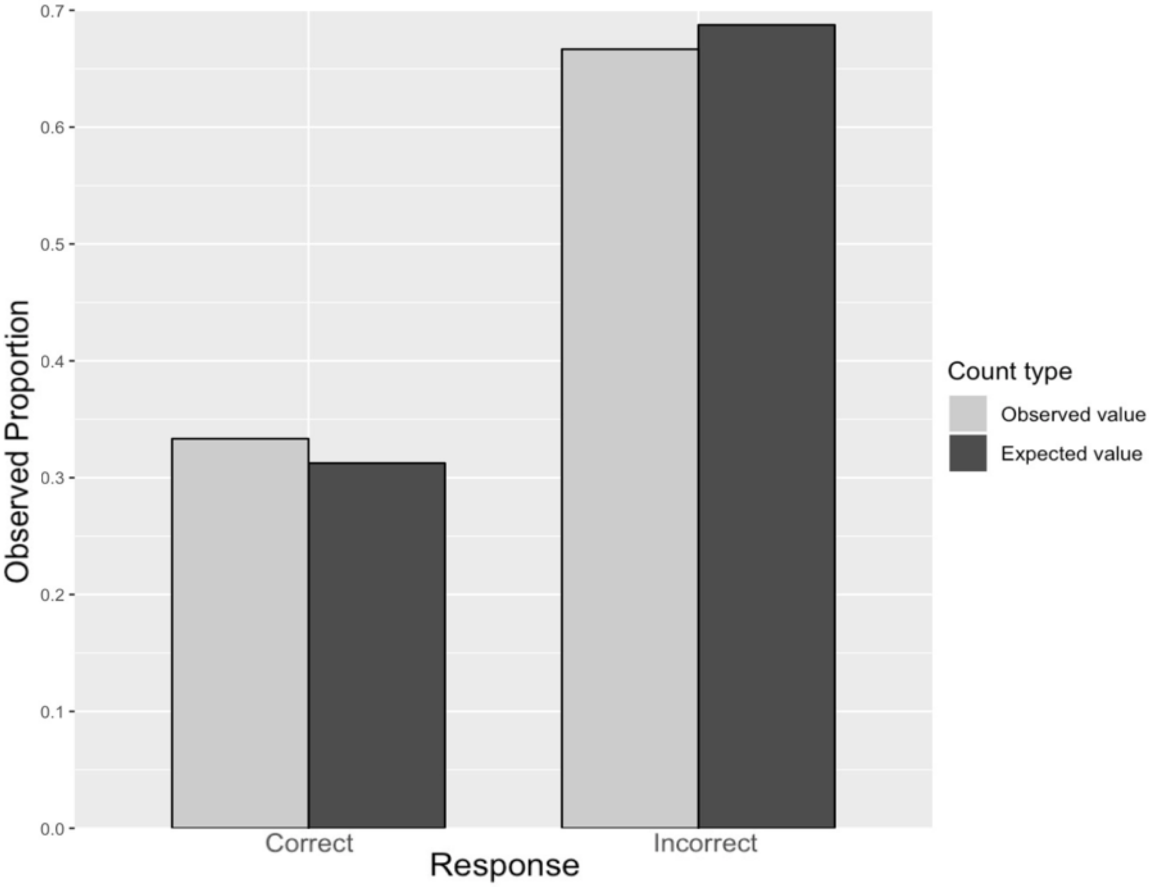


Figure 2. Level A responses: observed and expected outcomes based on chance. (*N* = 176)

predicting box exploration did not reveal any significance across any of the factors, but the GLMM model predicting box use did (see Tables 1 & 2). Male chimpanzees were predicted to be less likely to use the box (Estimate = -1.258, S.E. = 0.458, *z* value = -2.745, $p < 0.01$) and nonadult chimpanzees were predicted to be more likely to use the box (Estimate = 1.307, S.E. = 0.629, *z* value = 2.078, $p < 0.05$). The placement of the box in the porch area outside predicted a lower likelihood of box use (Estimate = -1.683, S.E. = 0.571, *z* value = -2.947, $p < 0.05$).

Table 1

Firth's Corrected Logistic Regression Model of All Predictors and Box Exploration

Fixed Effect	Coefficient	S.E.	Chisq	*p*
Males	-1.365	2.558	0.235	0.627
Nonadults	1.767	1.550	0.671	0.412
Sequence	-0.070	0.107	0.202	0.652
Up Position	1.009	1.072	0.445	0.504
Outside Location	1.360	1.223	0.798	0.371
Quarantine	0.893	1.504	0.225	0.634
Habitat Access	0.212	1.176	0.016	0.897
Original Program	-1.258	1.661	0.219	0.639

Table 2

GLMM of All Predictors and Box Use

Fixed Effect	Estimate	S.E.	*z* value	*p*
Males	-1.258	0.458	-2.745	0.006 **
Nonadults	1.307	0.629	2.078	0.037 *
Sequence	-0.028	0.049	-0.589	0.556
Up Position	0.105	0.464	0.226	0.821
Outside Location	-1.683	0.571	-2.947	0.003 **
Quarantine	-1.184	0.733	-1.614	0.106
Habitat Access	-0.831	0.512	-1.623	0.104
Original Program	0.631	0.691	0.914	0.360

Notes. * $p < 0.5$. **$p < 0.01$.

The model examining box interaction duration was significant for each fixed effect predictor (see Table 3). Box presentation sequence had a negative relationship with duration (Estimate = -0.10, S.E. = 0.003, *z* value = -32.81, $p < 0.001$). Chimpanzee sex also had a negative relationship for the male sex (Estimate = -2.238, S.E. = 0.059, *z* value

= -37.52, $p < 0.001$). Location of box placement was negative for outside placement (Estimate = -0.21, S.E. = 0.02, *z* value = -8.15, $p < 0.001$). Having access to the Peachtree habitat led to a negative relationship with duration of use (Estimate = -0.82, S.E. = 0.02, *z* value = -29.36, $p < 0.001$). The original program was negatively associated with duration (Estimate = -0.08, S.E. = 0.03, *z* value = -2.13, $p < 0.05$). Age class showed a positive relationship for the Nonadults class (Estimate = 0.87, S.E. = 0.04, *z* value = 20.64, $p < 0.001$). Box positioning at the upper level had a positive relation with duration (Estimate = 0.15, S.E. = 0.02, *z* value = 6.61, $p < 0.001$). Duration had a positive relationship with quarantine procedures (Estimate = 0.49, S.E. = 0.04, *z* value = 12.16, $p < 0.001$).

Table 3

GLMM of All Predictors and Duration

Fixed Effect	Estimate	S.E.	*z* value	*p*
Males	-2.238	0.059	-37.521	<0.0001 ***
Nonadults	0.876	0.042	20.644	<0.0001 ***
Sequence	-0.101	0.003	-32.817	<0.0001 ***
Up Position	0.157	0.023	6.619	<0.0001 ***
Outside Location	-0.211	0.025	-8.153	<0.0001 ***
Quarantine	0.496	0.040	12.169	<0.0001 ***
Habitat Access	-0.825	0.028	-29.364	<0.0001 ***
Original Program	-0.083	0.039	-2.139	0.0324 *

Notes. * $p < 0.5$. ** $p < 0.01$. *** $p < 0.001$.

The two GLMMs with box exploration and box use as the response variables and sequence as the sole predictor did not reveal significance for box use, but did reveal a significant negative relationship for box exploration (Estimate = -0.135, S.E. = 0.039, *z* value = -3.431, $p < 0.001$) (see Figure 3).

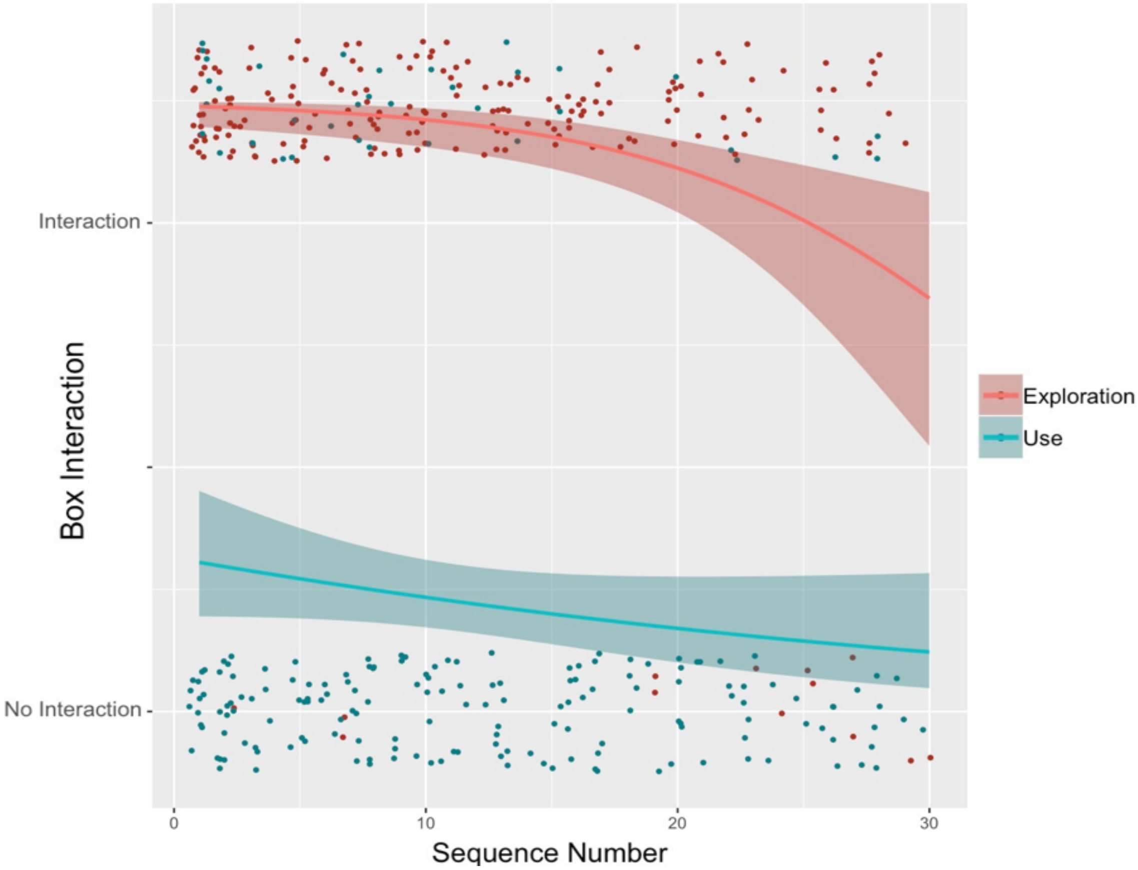


Figure 3. Exploration and use rates by sequence with mixed effect regression.

Stress Indicators and Enrichment Value

Behavioral data collected to assess possible stress induced by box interaction yielded 48 separate observations of SRB, with aggressive actions directed toward the box being the most frequently observed behavior (85%, $n = 41$) (Appendix B3). The other seven observations were of self-scratching during box interaction. One individual, Noel, accounted for 54% of total observed SRB. The results of the one-sample t-test of SRBs were significantly different than the expected mean of 0, t (194) = 4.04, $p < 0.0001$. Interaction duration as an indicator of enrichment value yielded results that were also significantly different from an anticipated mean of 0, $t\,(191) = 6.65$, $p < 0.0001$. (Appendix B4; DJT Females durations are from 1 hr of use only). Noel also interacted with the box for longer than any other chimpanzee, at 3,522 s (28%) of the total 12,234 s of all interaction time (Appendix B3).

CHAPTER V

DISCUSSION

Behavioral Observations

The behavioral observations I recorded should not be considered a comprehensive representation of the overall daily activity budget of the chimpanzees because of the narrow range during the day that they were collected. The chimpanzees were only observed between 11:30 am and 4:30 pm, which makes these observations susceptible to sampling bias (Altmann, 1974). However, this limitation was unavoidable, as this time range represented when caregivers could offer me access to areas to observe the focal chimpanzees. As an example of the potential skew of the data, aggressive behaviors did not account for even 1% of observed behaviors, but mornings in the villas were marked by high levels of excitement in the chimpanzees as they awaited their first meal, which is hand-served to each individual and, therefore, time-consuming to distribute. Prolonged excitement over the anticipation of meals sometimes led to displaying behavior and agonistic interactions between individuals, which was not reflected in my observations collected in the afternoon. In addition, behavioral observations began immediately before or after lunch was served as scattered forage, which may explain why the Feeding/Foraging category contained the second highest proportion of observed behaviors (15%); chimpanzees would spend a fair amount of time seeking out the individual pieces of scattered food around the villa. Lunch served as forage meant that the chimpanzees did not have to wait to obtain food and, as such, there was less excitement that culminated in agonistic interactions.

The high rate of inactivity observed (48%) may also be biased by the time of day during sampling. With the departure of the caregivers after providing lunch and afternoon enrichment, the chimpanzees were at liberty to occupy their afternoon however they chose, which I observed to be generally inactive. Bloomsmith and Lambeth (1995) observed high levels of inactivity between the meals of chimpanzees that were on predictable feeding schedules, which is consistent with the behavioral data I recorded for the afternoons. This inactivity is in stark contrast to the pre-meal excitement of breakfast and dinner, which is a common occurrence and known as food anticipatory behavior (Mistlberger, 1994). During behavioral observations and box presentations, I could confirm the arrival of the caregivers for dinner not only by the sound of their approaching vehicle but also by the directed alertness of the chimpanzees toward the approaching vehicle. The arrival of the vehicle could therefore have served as an unreliable signal of feeding events, which may have increased frustration and agonism when meals were not served promptly at arrival (Gottlieb, Coleman, & McCowan, 2013).

Another drawback of this sampling method was its bias for state behaviors over event behaviors. Behaviors included in the aggressive, affiliative, sexual, and abnormal categories are often rather brief, and 2-minute interval focal samples are not effective at capturing these quick moments well (Altmann, 1974). This explains the relatively low frequencies of the other behavioral categories outside of inactive and feeding/foraging categories. These two categories encompass behaviors that frequently last several minutes. The rate of abnormal behaviors is worth discussing as well, as the contentiousness of what behaviors can be considered abnormal means that the 5% frequency at which abnormal behaviors were observed could be interpreted as an inflated number if one eliminates some activities, such as coprophagy, from being labeled as abnormal (Hopper et al., 2016). Even at this inflated rate, the proportion of abnormal behavior is not as prevalent as frequencies observed in other populations (Birkett & Newton-Fisher, 2011). Despite the potential for a skewed representation of the activity budgets of the study subjects, this sampling method was beneficial in demonstrating that the chimpanzees were fairly inactive during the afternoon between lunch and dinner, thus indicating that the provision of stimulating enrichment during this time frame was appropriate.

Task Proficiency

Despite the numerous opportunities to interact with the box, the chimpanzees as a group did not perform above chance in the trials that they did complete. There are many explanations for this lack of success. First and foremost, as a voluntary task with no prior training on how to use the box's program, the chimpanzees' success was contingent upon not only their intelligence but also their curiosity and persistence (Litman, 2005). An individual would have needed to be sufficiently motivated to explore the box to the point that they discovered that the start button would initiate a trial. The individual would then need to continue their investigation by selecting a lit button rather than an unlit one. Next, the chimpanzees would have to stumble upon the correct response to the trial accidentally. This cycle of events would then need to be repeated many times for the chimpanzees to have the opportunity to discover the pattern of oddity equaling success. With all these steps acting as barriers impeding quick task comprehension, it is unsurprising that the chimpanzees as a whole were unable to respond above chance.

It is also important to mention that because only lit buttons were recorded as responses during a trial, there is the possibility that success rates were inflated as a consequence. Chimpanzees were sometimes observed to push buttons regardless of whether they were lit or not during a trial, and as was stated in the Methods section, unlit buttons were not responsive to button pushes and could not be recorded as incorrect inputs even though they technically were. The chimpanzees that initiated trials were sometimes observed to select buttons in both a clockwise and counterclockwise fashion by pushing adjacent buttons one after the other. This was not the case for each individual every time they used the box, but it occurred at least once among three individuals (AY, BC, HR). Some chimpanzees were also observed attempting to interact with the cues directly on the LCD screen (OC). These attempts imply that the box would have been more effective and intuitive had the screen and task been an actual touchscreen rather than button-dependent. The connection between the buttons and the cues on the screen may in fact have been counterintuitive and confusing for the participants and a major obstacle to interacting with the box. The matching of the white circle button to the white cue on display may have also added to the incomprehensibility of the task, in that it was in opposition to what the task was actually testing (i.e., nonmatching). Should these issues be the case, then the overall box design is flawed and should be reevaluated to improve intuitive use.

The number of completed trials (176) is a relatively low rate of box usage given the number of opportunities the chimpanzees had access to the box. This study was distinct from others involving interactive technology in that the chimpanzees did not receive any training on how to engage with the program, which is typical in other touchscreen studies (Herrelko et al., 2012; Calapai et al., 2017). As a visiting researcher, I did not have the opportunity or time to interact with the chimpanzees to the degree that would allow for training on box use. With the limited data set collected, it would be inadvisable to draw conclusions about the cognitive ability of chimpanzees in terms of inductive reasoning from these results.

Predictors of Box Interaction

The interaction frequencies for box exploration was higher at the beginning of the study as opposed to the end, but box use was low and consistently so. The chimpanzees' dwindling interest in the box suggests that they became habituated to its presence over time as its novelty waned (Murphy, McSweeney, Smith, & McComas, 2003). Efforts to rekindle interest in the box were made with program modifications to make the box more visually stimulating with the intermittently lit buttons and the addition of the simpler levels. Even with these changes, the chimpanzees' interest was not high enough to incite interaction with the box. In addition, box presentations for a villa were often intermittent, in that several days could go by before the chimpanzees would be given access to the box again, which is believed to bolster interest in enrichment (Murphy, McSweeney, Smith, & McComas, 2003) but did not appear to make a difference in this specific instance.

Sex was surprisingly significant in terms of predicting box usage. Males and females did not differ in their exploratory behaviors of the box, but males were less likely to use the box. This female bias echoes the discoveries made of wild chimpanzees and tool use for extractive foraging (Gruber, Clay, & Zuberbühler, 2010; McGrew, 1979; Pruetz & Bertolani, 2007). The tools used in termite-fishing and ant-dipping by the chimpanzees at Gombe were shown to be significantly higher among females than males, as evidenced by their longer extraction durations and higher insect content in their feces compared to males (McGrew, 1979). At Fongoli, chimpanzee spear-hunting of bushbabies as they sleep in tree hollows has shown to be a heavily female-centric activity and entails the precise processing of tree branches in order for them to function as adequate spears (Pruetz & Bertolani, 2007). In a sense, the box represents a kind of extractive foraging, which appears to be a preferred method of food acquisition for female chimpanzees both in the wild and in captivity.

Nonadult chimpanzees were predicted to be more likely to use the box. This may have been the result of a greater degree of curiosity and novelty-seeking in younger individuals, but no literature has found consistent age-related differences in neophilia (Kendal, Coe, & Laland, 2005). This nonadult positive relationship is likely the result of skew caused by one nonadult individual, Noel, who accounted for almost half of all box uses in conjunction with the low engagement of the Chimps Ahoy adult males. In addition, chimpanzees were predicted to be less likely to use the box when it was placed outdoors in the porch areas, but these results should be accepted tentatively because of the imbalance of box location presentations that occurred due to the onset of the winter season. This negative relationship could be the result of the correlation between the earlier sequence in which box presentations occurred in the outdoor location and may have skewed the results in that direction.

The model for duration showed significance for each predictor. The nonadult age class, elevated position, and quarantine predictors were positively correlated with duration, whereas sequence, the male sex, outside location, access to the habitat, and original program were associated with decreases in interaction duration. The estimates of the significant results from the box exploration GLMM are consistent with the estimates of the corresponding fixed effects in the duration of interaction GLMM (i.e., nonadults use the box more often/longer, males use the box less often/more briefly, and placing the box in the porch area makes the chimps less likely to use the box for long durations).

The behavioral data results suggest that the time of day that the boxes were presented may have been a key influence on box interaction rates as well. Because box presentations always followed the serving of lunch and afternoon enrichment, the chimpanzees may have been less motivated to seek out food when their preferred activity during this time was to be inactive or resting (Mistlberger, 1994; Bloomsmith & Lambeth,1995). Food served at lunch and as enrichment were much more readily accessible than the food pellets dispensed by the box. To compound the motivation problem, the food pellets themselves may not have been an enticing enough reward for the chimpanzees. Several chimpanzees were observed to drop their reward pellet immediately after tasting it, most notably Noel, who would actively throw the pellets through the mesh during the third box presentation at V4. Other chimpanzees were observed to consume these discarded rewards instead of the chimpanzee who had earned them.

Stress-Related Behaviors

The level of SRB observed indicates that the task was too challenging for the chimpanzees, which led to frustration and the subsequent self-directed behaviors and aggressions enacted toward the box. These rates of SRB and frustration may have also affected the rates of box interaction by outweighing the low motivation and value of the rewards. Noel expressed the most behaviors defined as SRB during her engagement with the box, and so we must decide whether its benefits in the form of increased activity is of greater value than the potential frustration it elicits (Meehan & Mench, 2007; Bennet et al., 2014; Dutton et al., 2018). Most of the aggressions consisted of attempts to pull or push the box from the mesh rather than hit it specifically and often occurred before use of the program. One interpretation of this may be that because the chimpanzees have never encountered this kind of enrichment before, and because a large part of their current enrichment entails some degree of destruction to acquire food rewards, that they believed the box was yet another form of enrichment that necessitated deconstruction to acquire the rewards.

In regard to Noel who accounted for over half of the observed SRB, it is important to establish context and know that she has a unique tendency toward destructive behaviors of many objects to which she is given access. This is exemplified by her removal of the plastic flaps covering the shifting doors separating the porches from the bedroom areas, and prying off of the plywood sheets used to create a barrier in the mesh between bedrooms when the new male group arrived in November. These behaviors might suggest that Noel is acting out of compromised welfare, but given her high degree of sociality and playfulness with chimpanzees and humans alike, this would be an unlikely assumption. Rather than conclude that Noel behaved aggressively toward the box as a result of frustration, her caregivers were more inclined to propose that her preferred enrichment is the exploration of the internal composition of objects; a testament to her inquisitive nature (Damerius et al., 2017). Another plausible explanation for her behavior could be neophobia, which is supported by her behavior during the first box presentation at V4 when she directed mild displaying behavior in the form of gentle swaying near the box after it emitted a chime for the first time. Previous studies have suggested that chimpanzees are the most cautious large-bodied ape when presented with novel food (Gustafsson, et al., 2014), and similar reluctance may have been experienced by the chimpanzees in this study.

Noel interacted with the box for the longest duration and was observed to perform a self-directed behavior only once during all her interactions, which may further the interpretation that her aggressive behaviors toward the box were not necessarily indicative of physiological arousal and stress-related behavior (Wagner, et al., 2016). Taking into account the unique dispositions of each individual is critical for a holistic interpretation of these results (Herrelko et al., 2012; Hopper et al., 2014). Because we are unable to ascertain the underlying motivations of the chimpanzees' actions, we must by default err on the side of caution and regard her aggression as an indicator of stress. Although tasks such as these may elicit indicators of stress in the animals that participate in them, it is important to recall that arguments can be made for some degree of stress being beneficial to overall welfare because it replicates the conditions these animals would experience in the wild (Meehan & Mench, 2007).

Enrichment Value

The duration of box interactions was significantly different than what the null hypothesis would predict (Appendix B4). The chimpanzees clearly expressed interest in the box, especially during the earlier presentations (see Figure 3). Between the significant results indicative of stress and the significant results of duration of use as enrichment value, it may seem difficult to decide whether the box's value as enrichment outweighs the stress it causes. Because both *t*-tests were similar in their degree of significance, we can look to the *t* value itself to aid our decision. For stress, the test yielded a *t* value equal to 4.04, whereas in the test for enrichment the *t* value was 6.65. Based on the magnitude of these numbers, the enrichment value of the box as measured by the duration of interactions is greater than the stress elicited by those interactions. Similar to Herrelko et al. (2012), repeated interest and engagement with the box suggests that the box was in fact enriching, although it may be more accurate to say that the box's novelty was the root cause of interest and enrichment. Dwindling interest does not necessarily indicate that the box was an inadequate form of enrichment, but rather that refinement of the box and how it is presented is required if we are to consider future box presentations with these chimpanzees. An alternative explanation for low box interaction may be that the chimpanzees were sufficiently stimulated by their ordinary day-to-day routine, such that the addition of the box was superfluous to their already enriched environment; meaning to say, Project Chimps fulfills its mission statement to provide exemplary care to the chimpanzees.

Future Directions

This study was designed by a novice to cognitive studies, and the experience generated many ideas to attempt the same study with better, more effective strategies. Some of the chimpanzees were inclined to directly touch the screen rather than the buttons, thus demonstrating that the use of touchscreens may be a critical element because they are more intuitive to use. The size of the box also proved to be unwieldy to attach to the mesh, but was necessary to accommodate both the larger screen and the buttons. Replacing both these components with a touchscreen would therefore reduce the size of the box and streamline its setup. Better food rewards would also be advantageous;

nuts and dried fruits are popular with the chimpanzees at Project Chimps and would have served well as incentive to participate in the task. This type of reward was precluded by the design of the gumball mechanism which was simple in its construction but ineffective at dispensing objects that were not spherical in shape. An alternative mechanism that can accommodate irregularly shaped items may make a significant difference.

Another redesign of the box that could have improved the study is screen sharing, which would allow the observer to see the display on the box as the chimpanzees saw it. The manner in which the box was hung on the mesh made viewing the display impossible in certain areas of the villa, so that the observer could not see what the cues on the screen were until reviewing the transcripts of the trials later. Such a feature necessitates strong Bluetooth and WiFi connections however, which was unavailable in the villas during the study. An alternative to remote screen sharing is integration of a second screen into the box itself on the opposite side of the one the chimpanzees have access to. This method would be straightforward to accomplish, and would not require a second touchscreen. Inclusion of software to record individual progress to tailor trial difficulty to that individual based on their level of success would also be a worthwhile addition, but was unnecessary for this study because no individual was able to progress past the first level. For future devices that are more engaging, an individual tailoring feature will be both essential and invaluable.

Training of individuals for box use may have been the catalyst necessary to facilitate chimpanzee success, but these practices may muddle results in that the subjects are not self-discovering solutions and are rather following a learned routine. Instead of classical operant conditioning training, designing simple tasks to precede cognitive tasks may act as an effective reinforcer for promoting interaction, especially when the device with which the individuals interact is first introduced. Once individuals demonstrate active participation, the cognitive task of interest can then be integrated into the device. Additionally, as a visiting researcher, I had neither the authority nor the time to commit to training the potential subjects. This study may have yielded different results if executed by an individual who is allowed greater interaction with the chimpanzees.

Many cognitive studies approach assessments of cognition from only one angle by testing for a single cognitive ability, as in this study. Creating a computerized enrichment device that can offer an array of tasks and activities would be a more effective and productive method of cognitive assessment and enrichment by adding agency to which tasks a captive animal can participate in. Cognitive studies and captive enrichment need not be mutually exclusive; all that is required to join these two pursuits is imagination and creativity.

Conclusions

No evidence was discovered to suggest that the chimpanzees could reason inductively. However, absence of evidence is not evidence of absence. Little data were gathered from which we could investigate this ability. Future, more robust studies may yet demonstrate this ability. This study did reveal a decrease in interaction over time consistent with habituation and a negative tendency for males to use the program, which echoes their lack of engagement in extractive foraging in the wild. Nonadults were also found to be more likely to engage with the box. The lengths of duration that the chimpanzees interacted with the box supports its provision as enrichment and, although it may have incited some SRBs, there is reason to believe that the stress rates are inflated because of one subject’s propensity toward destructive behavior and that the enrichment value outweighs the stress regardless. The negative results in the cognitive aspect of analysis may discourage attempts at future investigations, but in the words of de Waal (2010, p. 303) “negative experimental outcomes have more often than not been followed by positive ones after modification of the experimental design.”

REFERENCES

Allen, C. & Trestman, M. (2017). “Animal Consciousness”, *The Stanford Encyclopedia of Philosophy.* Retrieved from https://plato.stanford.edu/entries/consciousness-animal/#hist.

Allgas, N., Shanee, S., Shanee, N., Chambers, J., Tello-Alvarado, J. C., Keeley, K., & Pinasco, K. (2017). Natural re-establishment of a population of a critically endangered primate in a secondary forest: The San Martin titi monkey (*Plecturocebus oenanthe*) at the Pucunucho Private Conservation Area, Peru. *Primates*, *58*(2), 335-342.

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*(3), 227-266.

Altschul, D. M., Wallace, E. K., Sonnweber, R., Tomonaga, M., & Weiss, A. (2017). Chimpanzee intellect: Personality, performance and motivation with touchscreen tasks. *Royal Society Open Science*, *4*(5), 170169.

Ash, I., Jee, B., & Wiley, J. (2012). Investigating insight as sudden learning. *The Journal of Problem Solving, 4*(2), 1-27.

Baker, K., & Aureli, F. (1997). Behavioural indicators of anxiety: An empirical test in chimpanzees. *Behaviour*, *134*(13), 1031-1050.

Baker, K., Weed, J., Crockett, C., & Bloomsmith, M. (2007). Survey of environmental enhancement programs for laboratory primates. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, *69*(4), 377-394.

Baker, K. C. (2016). Survey of 2014 behavioral management programs for laboratory primates in the United States. *American Journal of Primatology*, *78*(7), 780-796.

Banerjee, K., Chabris, C., Johnson, V., Lee, J., Tsao, F., Hauser, M., & Ferrari, P. (2009). General intelligence in another primate: Individual differences across cognitive task performance in a new world monkey (*Saguinus oedipus*). *PLoS ONE, 4*(6), E5883.

Bassett, L., & Buchanan-Smith, H. M. (2007). Effects of predictability on the welfare of captive animals. *Applied Animal Behaviour Science*, *102*(3-4), 223-245.

Bear, M., Connors, B., & Paradiso, M. (2007). *Neuroscience: Exploring the brain* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Bennett, A., Perkins, C., Harty, N., Niu, M., Buelo, A., Luck, M., & Pierre, P. (2014). Assessment of foraging devices as a model for decision-making in nonhuman primate environmental enrichment. *Journal of the American Association for Laboratory Animal Science*, *53*(5), 452-463.

Bennett, A., Perkins, C., Tenpas, P., Reinebach, A., & Pierre, P. (2016). Moving evidence into practice: Cost analysis and assessment of macaques' sustained behavioral engagement with videogames and foraging devices. *American Journal of Primatology*, *78*(12), 1250-1264.

Bernstein, I., & Webster, M. (1980). Learned taste aversions in humans. *Physiology & Behavior,25*(3), 363-366.

Birkett, L., & Newton-Fisher, N. (2011). How abnormal is the behaviour of captive, zoo-living chimpanzees? *PLoS One, 6*(6), E20101.

Bloomsmith, M., Brent, L., & Schapiro, S. (1991). Guidelines for developing and managing an environmental enrichment program for nonhuman primates. *Laboratory Animal Science, 41*(4), 372-7.

Bloomsmith, M., & Lambeth, S. (2000). Videotapes as enrichment for captive chimpanzees (*Pan troglodytes*). *Zoo Biology*, *19*(6), 541-551.

Boesch, C. (2007). What makes us human (*Homo sapiens*)? The challenge of cognitive cross-species comparison. *Journal of Comparative Psychology*, *121*(3), 227.

Bond, A., Kamil, A., & Balda, R. (2003). Social complexity and transitive inference in corvids. *Animal Behaviour, 65*(3), 479-487.

Bonnie, K., Milstein, M., Calcutt, S., Ross, S., Wagner, K., & Lonsdorf, E. (2012). Flexibility and persistence of chimpanzee (*Pan Troglodytes*) foraging behavior in a captive environment. *American Journal of Primatology, 74*(7), 661-668.

Bräuer, J., Kaminski, J., Riedel, J., Call, J., Tomasello, M., & Burghardt, G.M. (2006). Making inferences about the location of hidden food: Social dog, causal ape. *Journal of Comparative Psychology, 120*(1), 38-47.

Broom, D. (1991). Animal welfare: Concepts and measurement. *Journal of Animal Science, 69*(10), 4167-75.

Byrne, R. (2000). Evolution of primate cognition. *Cognitive Science, 24*(3), 543-570.

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., & Gail, A. (2017). A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. *Behavior Research Methods*, *49*(1), 35-45.

Call, J., Hare, B., & Tomasello, M. (1998). Chimpanzee gaze following in an object-choice task. *Animal Cognition, 1*(2), 89-99.

Call, J. (2000). Representing space and objects in monkeys and apes. *Cognitive Science*, *24*(3), 397-422.

Call, J. (2006). Inferences by exclusion in the great apes: The effect of age and species. *Animal Cognition, 9*(4), 393-403.

Call, J. & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187-192.

Carenzi, C., & Verga, M. (2009). Animal welfare: Review of the scientific concept and definition. *Italian Journal of Animal Science, 8*, 21-30.

Chalmeau, R., Lardeux, K., Brandibas, P., & Gallo, A. (1997). Cooperative problem solving by orangutans (*Pongo pygmaeus*). *International Journal of Primatology, 18*(1), 23-32.

Chalmeau, R., Visalberghi, E., & Gallo, A. (1997). Capuchin monkeys, *Cebus apella* fail to understand a cooperative task. *Animal Behaviour, 54*(5), 1215-25.

Cheney, D., & Seyfarth, R. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.

Chrobak, J., Hanin, I., Lorens, S., Napier, T., & Squire, L. (1995). Within-subject decline in delayed-non-match-to-sample radial arm maze performance in aging Sprague-Dawley rats. *Behavioral Neuroscience, 109*(2), 241-245.

Clark, F. (2011). Great ape cognition and captive care: Can cognitive challenges enhance well-being? *Applied Animal Behaviour Science, 135*(1), 1-12.

Clark, F., & Smith, L. (2013). Effect of a cognitive challenge device containing food and non-food rewards on chimpanzee well-being. *American Journal of Primatology*, *75*(8), 807-816.

Clay, A., Perdue, B., Gallema, D., Dolins, F., & Bloomsmith, M. (2011). The use of technology to enhance zoological parks. *Zoo Biology*, *30*(5), 487-497.

Cohen, J. (2007). Relative differences: the myth of 1%. *Science*, *316*(5833), 1836.

Crofoot, M. C. (2013). The cost of defeat: Capuchin groups travel further, faster and later after losing conflicts with neighbors. *American Journal of Physical Anthropology*, *152*(1), 79-85.

Csatádi, K., Leus, K., & Pereboom, J. (2008). A brief note on the effects of novel enrichment on an unwanted behaviour of captive bonobos. *Applied Animal Behaviour Science, 112*(1), 201-204.

Damerius, L.A., Graber, S.M., Willems, E.P., & van Schaik, C.P. (2017). Curiosity boosts orang-utan problem-solving ability. *Animal Behaviour, 134*, 57-70.

Decasien, A., Williams, S., & Higham, J. (2017). Primate brain size is predicted by diet but not sociality. *Nature Ecology & Evolution, 1*(5), 112.

De Waal, F.B. (2010). The need for a bottom-up approach to chimpanzee cognition. In E. Lonsdorf, S. Ross, T. Matsuzawa, & J. Goodall. *The mind of the chimpanzee: Ecological and experimental perspectives* (pp. 296-306). Chicago: University of Chicago Press.

Dunbar, R., & Shultz, S. (2007). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1480), 649-58.

Dutton, M. B., Pierre, P. J., Bailoo, J., Warkins, E., Michel, G. F., & Bennett, A. J. (2018). A model quantitative assessment tool for nonhuman primate environmental enrichment plans. *BioRxiv*, 341206.

Egelkamp, C. L., & Ross, S. R. (2019). A review of zoo-based cognitive research using touchscreen interfaces. *Zoo biology*, *38*(2), 220-235.

Elder, C., & Menzel, C. (2001). Dissociation of cortisol and behavioral indicators of stress in an orangutan (*Pongo pygmaeus*) during a computerized task. *Primates*, *42*(4), 345-357.

Elliott, R., Norris, C., Ettlinger, E., & Mishkin, G. (1977). Some factors influencing nonmatching to sample in the monkey. *Bulletin of the Psychonomic Society, 9*(6), 395-396.

Erdőhegyi, Á., Topál, J., Virányi, Z., & Miklósi, Á. (2007). Dog-logic: Inferential reasoning in a two-way choice task and its restricted use. *Animal Behaviour, 74*(4), 725-737.

Fagot, J., & Maugard, A. (2013). Analogical reasoning in baboons (*Papio papio*): Flexible reencoding of the source relation depending on the target relation. *Learning & Behavior, 41*(3), 229-237.

Falcone, R., Bevacqua, S., Cerasti, E., Brunamonti, E., Cervelloni, M., & Genovesio, A. (2013). Transfer of the nonmatch-to-goal rule in monkeys across cognitive domains. *PloS One*, 8(12), e84100.

Flemming, T., Beran, M., Thompson, R., Kleider, H., Washburn, D., & Burghardt, Gordon M. (2008). What meaning means for same and different: Analogical reasoning in humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology, 122*(2), 176-185.

Forss, S. I., Schuppli, C., Haiden, D., Zweifel, N., & van Schaik, C. P. (2015). Contrasting responses to novelty by wild and captive orangutans. *American Journal of Primatology*, *77*(10), 1109-1121.

Garber, P., & Dolins, F. (2014). Primate spatial strategies and cognition: Introduction to this special issue. *American Journal of Primatology, 76*(5), 393-398.

Gillan, D.J., Premack, D., Woodruff, G., & Wagner, A. (1981). Reasoning in the chimpanzee: I. Analogical reasoning. *Journal of Experimental Psychology: Animal Behavior Processes, 7*(1), 1-17.

Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*(2), 150.

Goodman, M., Porter, C., Czelusniak, J., Page, S., Schneider, H., Shoshani, J., Gunnell, G., & Groves, C. (1998). Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. *Molecular Phylogenetics and Evolution, 9*(3), 585-598.

Gottlieb, D. H., Coleman, K., & McCowan, B. (2013). The effects of predictability in daily husbandry routines on captive rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science*, *143*(2-4), 117-127.

Grimm, D. (June 12, 2015). Research chimps to be listed as "endangered". *Science*. Retrieved from https://www.sciencemag.org/news/2015/06/research-chimps-be-listed-endangered?_ga=1.147171703.1766306986.1392323490.

Gronqvist, G., Kingston-Jones, M., May, A., & Lehmann, J. (2013). The effects of three types of environmental enrichment on the behaviour of captive Javan gibbons (*Hylobates moloch*). *Applied Animal Behaviour Science, 147*(1-2), 214-223.

Gruber, T., Clay, Z., & Zuberbühler, K. (2010). A comparison of bonobo and chimpanzee tool use: evidence for a female bias in the Pan lineage. *Animal Behaviour*, *80*(6), 1023-1033.

Grunauer, P. P., & Walguarnery, J. W. (2018). Relative response to digital tablet devices and painting as sensory enrichment in captive chimpanzees. *Zoo Biology*, *37*(4), 269-273.

Gustafsson, E., Saint Jalme, M., Bomsel, M. C., & Krief, S. (2014). Food neophobia and social learning opportunities in great apes. *International Journal of Primatology*, *35*(5), 1037-1071.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(1), 139-151.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 771-785.

Hare, B., Melis, A., Woods, V., Hastings, S., & Wrangham, R. (2007). Tolerance allows bonobos to outperform chimpanzees on a cooperative task. *Current Biology, 17*(7), 619-623.

Heberlein, M., Turner, D., Range, F. & Virányi, Z. (2016). A comparison between wolves, *Canis lupus*, and dogs, *Canis familiaris*, in showing behaviour towards humans. *Animal Behaviour, 122*, 59-66.

Hernández-Pacheco, R., Rawlins, R. G., Kessler, M. J., Williams, L. E., Ruiz-Maldonado, T. M., González-Martínez, J., ... & Sabat, A. M. (2013). Demographic variability and density-dependent dynamics of a free-ranging rhesus macaque population. *American Journal of Primatology*, *75*(12), 1152-1164.

Herrelko, E. S., Vick, S. J., & Buchanan-Smith, H. M. (2012). Cognitive research in zoo-housed chimpanzees: Influence of personality and impact on welfare. *American Journal of Primatology*, *74*(9), 828-840.

Herrmann, E., Call, J., Hernàndez-Lloreda, M., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science, 317*(5843), 1360-6.

Hill, A., Collier-Baker, E., Suddendorf, T., & Burghardt, G. M. (2011). Inferential reasoning by exclusion in great apes, lesser apes, and spider monkeys. *Journal of Comparative Psychology, 125*(1), 91-103.

Hill, A., Collier-Baker, E., Suddendorf, T., & Call, J. (2012). Inferential reasoning by exclusion in children (*Homo sapiens*). *Journal of Comparative Psychology, 126*(3), 243-254.

Hockings, K. J., Yamakoshi, G., Kabasawa, A., & Matsuzawa, T. (2010). Attacks on local persons by chimpanzees in Bossou, Republic of Guinea: Long-term perspectives. *American Journal of Primatology*, *72*(10), 887-896.

Hook, M., Lambeth, S., Perlman, J., Stavisky, R., Bloomsmith, M., & Schapiro, S. (2002). Inter-group variation in abnormal behavior in chimpanzees (*Pan troglodytes*) and rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science,76*(2), 165-176.

Hopper, L.M., Price, M., Freeman, S., Lambeth, A., Schapiro, H., & Kendal, D. (2014). Influence of personality, age, sex, and estrous state on chimpanzee problem-solving success. *Animal Cognition, 17*(4), 835-847.

Hopper, L.M., Kurtycz, L., Ross, S., Bonnie, K., & Vonk, J. (2015). Captive chimpanzee foraging in a social setting: A test of problem solving, flexibility, and spatial discounting. *PeerJ, 3*, e833.

Hopper, L. M., Freeman, H. D., & Ross, S. R. (2016). Reconsidering coprophagy as an indicator of negative welfare for captive chimpanzees. *Applied Animal Behaviour Science*, *176*, 112-119.

Hulac, D., Benson, N., Nesmith, M. C., & Wollersheim Shervey, S. (2016). Using variable interval reinforcement schedules to support students in the classroom: An introduction with illustrative examples. *Journal of Educational Research and Practice*, *6*(1), 6.

Ikeda, Y. (2009). A perspective on the study of cognition and sociality of cephalopod mollusks, a group of intelligent marine invertebrates. *Japanese Psychological Research, 51*(3), 146-153.

Jacobson, S., Ross, S., & Bloomsmith, M. (2016). Characterizing abnormal behavior in a large population of zoo-housed chimpanzees: Prevalence and potential influencing factors. *PeerJ*, *4*, e2225.

Jelbert, S., Taylor, A., Gray, R., & Call, J. (2015). Reasoning by exclusion in New Caledonian crows (*Corvus moneduloides*) cannot be explained by avoidance of empty containers. *Journal of Comparative Psychology, 129*(3), 283-290.

Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science (New York, N.Y.), 153*(3735), 501-6.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition, 109*(2), 224-234.

Kaneko, T., & Tomonaga, M. (2011). The perception of self-agency in chimpanzees (*Pan troglodytes*). *Proceedings of the Royal Society B: Biological Sciences*, *278*(1725), 3694- 3702.

Kendal, R. L., Coe, R. L., & Laland, K. N. (2005). Age differences in neophilia, exploration, and innovation in family groups of callitrichid monkeys. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, *66*(2), 167-188.

Kim-Mccormack, N., Smith, C. & Behie, A. (2016). Is interactive technology a relevant and effective enrichment for captive great apes? *Applied Animal Behaviour Science, 185*, 1-8.

Kittler, K., Schnoell, A., & Fichtel, C. (2015). Cognition in ring-tailed lemurs. *Folia Primatologica; International Journal of Primatology, 86*(1-2), 106-16.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: Comparative studies on external morphology of the primate eye. *Journal of Human Evolution,40*(5), 419-35.

Koolhaas, J. M., Korte, S. M., De Boer, S. F., Van Der Vegt, B. J., Van Reenen, C. G., Hopster, H., ... & Blokhuis, H. J. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neuroscience & Biobehavioral Reviews*, *23*(7), 925-935.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science (New York, N.Y.), 354*(6308), 110-114.

Leavens, D., Aureli, F., Hopkins, W., & Hyatt, C. (2001). Effects of cognitive challenge on self-directed behaviors by chimpanzees (*Pan troglodytes*). *American Journal of Primatology: Official Journal of the American Society of Primatologists*, *55*(1), 1-14.

Lee, P. C. (2010). Sharing space: can ethnoprimatology contribute to the survival of nonhuman primates in human-dominated globalized landscapes?. *American Journal of Primatology*, *72*(10), 925-931.

Limongelli, L., Boysen, S., Visalberghi, E., & Snowdon, C.T. (1995). Comprehension of cause-effect relations in a tool-using task by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology,109*(1), 18-26.

Litman, J. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & emotion*, *19*(6), 793-814.

Lopresti-Goodman, S. M., Bezner, J., & Ritter, C. (2015). Psychological distress in chimpanzees rescued from laboratories. *Journal of Trauma & Dissociation*, *16*(4), 349-366.

MacDonald, S., Wilkie, D., & Gallup, G. (1990). Yellow-nosed monkeys' (*Cercopithecus ascanius whitesidei*) spatial memory in a simulated foraging environment. *Journal of Comparative Psychology, 104*(4), 382-387.

MacLean, E., Merritt, D., & Brannon, E. (2008). Social complexity predicts transitive reasoning in prosimian primates. *Animal Behaviour, 76*(2), 479-486.

Mallavarapu, S., Bloomsmith, M. A., Kuhar, C. W., & Maple, T. L. (2013). Using multiple joystick systems in computerised enrichment for captive orangutans. *Animal Welfare*, *22*(3), 401-409.

Marino, L., Connor, R., Fordyce, R., Herman, L., Hof, P., Lefebvre, …, & Whitehead, H. (2007). Cetaceans have complex brains for complex cognition. *PLoS Biology, 5*(5), E139.

Marsh, H., & MacDonald, L. (2012). Information seeking by orangutans: A generalized search strategy? *Animal Cognition, 15*(3), 293-304.

Marshall-Pescini, S., Schwarz, J., Kostelnik, I., Virányi, Z., & Range, F. (2017). Importance of a species' socioecology: Wolves outperform dogs in a conspecific cooperation task. *Proceedings of the National Academy of Sciences of the United States of America, 114*(44), 11793-11798.

Mason, G., & Latham, N. (2004). Can't stop, won't stop: Is stereotypy a reliable animal welfare indicator? *Animal Welfare, 13*(SS), S57-S69.

Matsuzawa, T. (2003). The Ai project: Historical and ecological contexts. *Animal cognition*, *6*(4), 199-211.

Matzel, L. D., & Sauce, B. (2017). Individual differences: Case studies of rodent and primate intelligence. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(4), 325.

McGrew, W. S. (1979). Evolutionary implications of sex differences in chimpanzee predation and tool use. In D. Hamburg & E. McCown (pp. 441-463). *The great apes*. Menlo Park, CA: Benjamin/Cummings Publishing Company.

Meehan, C., & Mench, J. (2007). The challenge of challenge: Can problem solving opportunities enhance animal welfare? *Applied Animal Behaviour Science*, *102*(3-4), 246-261.

Mistlberger, R. E. (1994). Circadian food-anticipatory activity: Formal models and physiological mechanisms. *Neuroscience & Biobehavioral Reviews*, *18*(2), 171-195.

Morgan, K. N., & Tromborg, C. T. (2007). Sources of stress in captivity. *Applied Animal Behaviour Science*, *102*(3-4), 262-302.

Müller, J., Massen, J., Bugnyar, T., & Osvath, M. (2017). Ravens remember the nature of a single reciprocal interaction sequence over 2 days and even after a month. *Animal Behaviour,128*, 69-78.

Murphy, E. S., McSweeney, F. K., Smith, R. G., & McComas, J. J. (2003). Dynamic changes in reinforcer effectiveness: Theoretical, methodological, and practical implications for applied research. *Journal of Applied Behavior Analysis*, *36*(4), 421-438.

Nawroth, C., von Borell, E., & Langbein, J. (2014). Exclusion performance in dwarf goats (*Capra aegagrus hircus*) and sheep (*Ovis orientalis aries*). *PLoS One*, *9*(4), e93534.

Neal Webb, S., Hau, J., & Schapiro, S. (2018). Captive chimpanzee (*Pan troglodytes*) behavior as a function of space per animal and enclosure type. *American Journal of Primatology, 80*(3), e22749.

Neldner, K., Collier-Baker, E., & Nielsen, M. (2015). Chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*) know when they are ignorant about the location of food. *Animal Cognition, 18*(3), 683-699.

Nishida, T., Corp, N., Hamai, M., Hasegawa, T., Hiraiwa-Hasegawa, M., Hosaka, K., … Zamma, K. (2003). Demography, female life history, and reproductive profiles among the chimpanzees of Mahale. *American Journal of Primatology*, *59*(3), 99-121.

Osvath, M. & Osvath, H. (2008). Chimpanzee (*Pan troglodytes*) and orangutan (*Pongo abelii*) forethought: Self-control and pre-experience in the face of future tool use. *Animal Cognition, 11*(4), 661-674.

Pacelle, W. (April 22, 2016). A year after New York Blood Center abandoned them, Liberian chimps are thriving under humane care. *A Humane World: Kitty Block's Blog*. Retrieved from https://blog.humanesociety.org/2016/04/liberian-chimpanzees-thrive-under-hsus-care.html?credit=blog_post_id8101.

Pepperberg, I., Koepke, A., Livingston, P., Girard, M., Ann Hartsfield, L., & Call, J. (2013). Reasoning by inference: Further studies on exclusion in grey parrots (*Psittacus erithacus*). *Journal of Comparative Psychology, 127*(3), 272-281.

Perdue, B., Clay, A., Gaalema, D., Maple, T., & Stoinski, T. (2012). Technology at the zoo: The influence of a touchscreen computer on orangutans and zoo visitors. *Zoo Biology, 31*(1), 27-39.

Platt, M. L., Seyfarth, R. M., & Cheney, D. L. (2016). Adaptations for social cognition in the primate brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1687), 20150096.

Plotnik, J., Lair, R., Suphachoksahakun, W., & de Waal, F. (2011). Elephants know when they need a helping trunk in a cooperative task. *Proceedings of the National Academy of Sciences of the United States of America, 108*(12), 5116-5121.

Pomerantz, O., & Terkel, J. (2009). Effects of positive reinforcement training techniques on the psychological welfare of zoo-housed chimpanzees (*Pan troglodytes)*. *American Journal of Primatology, 71*(8), 687-695.

Premack, D. & Premack, A. (1994). Levels of causal understanding in chimpanzees and children. *Cognition, 50*, 347-362.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526.

Pruetz, J. D., & Bertolani, P. (2007). Savanna chimpanzees, *Pan troglodytes verus*, hunt with tools. *Current biology*, *17*(5), 412-417.

R Core Team. (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. https://www.R-project.org/.

Radick, G. (2007). *The Simian Tongue: The long debate about animal language*. Chicago: University of Chicago Press.

Ramsey, G., Bastian, M. L., & van Schaik, C. (2007). Animal innovation defined and operationalized. *Behavioral and Brain Sciences*, *30*(4), 393-407.

Reader, S., & Laland, K. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences of the United States of America, 99*(7), 4436-4441.

Reader, S., Hager, Y., & Laland, K. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 366*(1567), 1017-27.

Ross, S. (2010). How cognitive studies help shape our obligation for the ethical care of chimpanzees. In E. Lonsdorf, S. Ross, T. Matsuzawa, & J. Goodall. *The mind of the chimpanzee: Ecological and experimental perspectives* (pp. 309-319). Chicago: University of Chicago Press.

Schmelz, M., Call, J., & Tomasello, M. (2011). Chimpanzees know that others make inferences. *Proceedings of the National Academy of Sciences*, *108*(7), 3077-3079.

Schmelz, M., & Call, J. (2016). The psychology of primate cooperation and competition: A call for realigning research agendas. *Phil. Trans. R. Soc. B*, *371*(1686), 20150067.

Schmitt, V., Pankau, B., & Fischer, J. (2012). Old world monkeys compare to apes in the primate cognition test battery. *PloS One, 7*(4), E32024.

Schneider, A., Melis, A., & Tomasello, M. (2012). How chimpanzees solve collective action problems. *Proceedings. Biological Sciences, 279*(1749), 4946-54.

Shettleworth, S. (2001). Animal cognition and animal behaviour. *Animal Behaviour*, *61*(2), 277-286.

Stone, A., Battistuzzi, F., Kubatko, L., Perry, G., Trudeau, E., Lin, H., & Kumar, S. (2010). More reliable estimates of divergence times in *Pan* using complete mtDNA sequences and accounting for population structure. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 365*(1556), 3277-88.

Taniuchi, T., Miyazaki, R., & Siddik, M. (2017). Concurrent learning of multiple oddity discrimination in rats. *Behavioural Processes, 140*, 6-15.

Taylor, A. (2014). Corvid cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(3), 361-372.

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

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends in Cognitive Sciences, 7*(4), 153-156.

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution, 52*(3), 314-320.

Tomonaga, M., Imura, T., & Emery, N. (2010). Visual search for human gaze direction by a chimpanzee (*Pan troglodytes*). *PLoS ONE, 5*(2), E9131.

Tulving, E. (1972). Organisation of Memory. New York: Academic Press.

Vaesen, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences, 35*(4), 203-218.

Van Leeuwen, E. J., Cronin, K. A., & Haun, D. B. (2014). A group-specific arbitrary tradition in chimpanzees (*Pan troglodytes*). *Animal Cognition*, *17*(6), 1421-1425.

Van Schaik, C. P., & Burkart, J. M. (2011). Social learning and evolution: The cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1008-1016.

Vasconcelos, M. (2008). Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes,78*(3), 313-334.

Visalberghi, E., Limongelli, L., & Gallup, G. (1994). Lack of comprehension of cause–effect relations in tool-using capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 108*(1), 15-22.

Visalberghi, E. (1997). Success and understanding in cognitive tasks: A comparison between *Cebus apella* and *Pan troglodytes*. *International Journal of Primatology, 18*(5), 811-830.

Wagner, K., Hopper, L., & Ross, S. (2016). Asymmetries in the production of self-directed behavior by chimpanzees and gorillas during a computerized cognitive test. *Animal Cognition, 19*(2), 343-350.

Walsh, S., Bramblett, C., & Alford, P. (1982). A vocabulary of abnormal behaviors in restrictively reared chimpanzees. *American Journal of Primatology*, *3*(1-4), 315-319.

Washburn, D., & Rumbaugh, D. (1992). Comparative assessment of psychomotor performance: Target prediction by humans and macaques (*Macaca mulatta*). *Journal of Experimental Psychology: General*, *121*(3), 305.

Washburn, D. (2015). The four Cs of psychological wellbeing: Lessons from three decades of computer-based environmental enrichment. *Animal Behavior and Cognition*, *2*(3), 218- 232.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655-684.

Whiten, A., Goodall, J., Mcgrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C.E.G., Wrangham, R.W., & Boesch, C. (1999). Cultures in chimpanzees. *Nature, 399*(6737), 682-685.

Yamanashi, Y. & Hayashi, M. (2011). Assessing the effects of cognitive experiments on the welfare of captive chimpanzees (*Pan troglodytes*) by direct comparison of activity budget between wild and captive chimpanzees. *American Journal of Primatology, 73*(12), 1231-1238.

APPENDIX

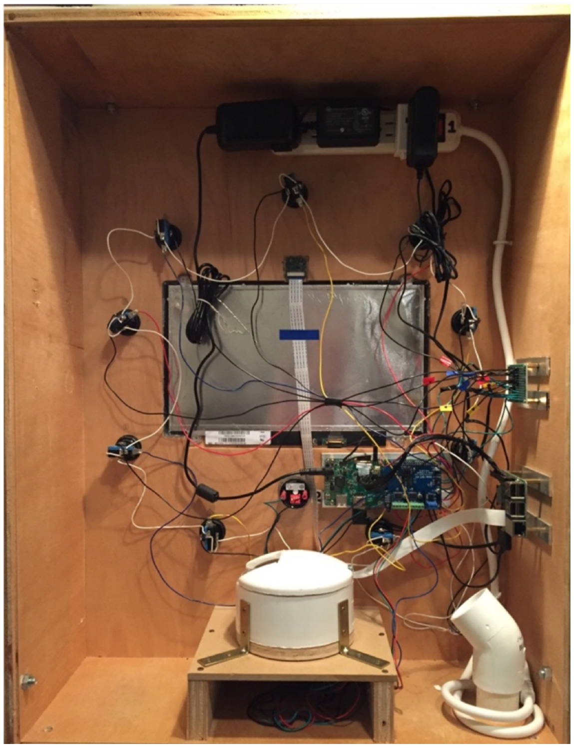
APPENDIX A

PHOTOS OF SANCTUARY AND BOXES

A1. Layout of habitat facilities at Project Chimps. Source: projectchimps.org

A2. Porch enclosures of the chimpanzee villas (without patio fencing). Source: projectchimp.org

A3. The display cues before a trial (a, left) and during Level D (b, right).



A4. Internal components of a box.

APPENDIX B

SUPPLEMENTARY TABLES AND FIGURE

B1

Demographics of Study Population

Individual	ID	Villa	Sex	Age	Behavioral Observations
Buttercup	BC	DJT	F	13	6
Charisse	CS	DJT	F	13	6
Emma	EM	DJT	F	13	0
Binah	BI	DJT	M	10	5
Danner	DR	DJT	M	8	6
Hercules	HC	DJT	M	11	6
Jacob	JC	DJT	M	7	5
Kennedy	KD	DJT	M	10	5
Kivuli	KU	DJT	M	7	1
Leo	LE	DJT	M	11	6
Oscar	OC	DJT	M	7	6
Ray	RY	DJT	M	9	4
Arthur	AU	CA	M	16	5
Josh	JH	CA	M	27	3
Marlon	MR	CA	M	17	6
Patrick	PT	CA	M	16	6
Quintin	QU	CA	M	16	3
Taz	TZ	CA	M	22	6
Almasi	AL	V4	F	11	6
Amy	AY	V4	F	9	4
Babs	BB	V4	F	22	5
Harley	HR	V4	F	9	6
LB	LB	V4	F	21	5
Loretta	LO	V4	F	9	6
Noel	NO	V4	F	11	4
Sarah	SR	V4	F	9	6
Sky	SY	V4	F	21	4

B2

Ethogram of Behaviors

Category	Code	Behavior	Description
Aggressive	AL	Agitated locomotion	Brisk, or rapid walking that often occurs with increased vigilance toward event, animal, or object.
	HA	Hit away	Source of apprehension or tension (another animal or object) is hit towards. If this occurs with screaming or fear grimace, will be recorded as such using modifiers.
	DS	Display	Aggressive behavior without any clear and identifiable recipient. May include pilo-erection, and such behaviors as beating on or moving inanimate objects, stomping, slapping, swaying, hooting, chest-beat, or running.
	TH	Threat	Includes arm raising, arm waving, wrist shaking, lunging toward, and chasing toward another animal. No physical contact is made.
	CA	Contact aggression	Aggressive behavior that involves physical contact between individuals. Includes, wrestling, lunge hit, grab, bite, and scratch.
Locomotion	HN	Hang	All of the animal's weight is supported by wire of walls (i.e. animal is grasping wire with hands and feet), or the animal is hanging beneath the climbing bars.
	WK	Walk	Moving through space at a calm, steady pace on horizontal surface (may be on ground, stairs, or platform). If walking occurs with play face, then it is considered play. Includes walking with food.

	CL	Climb	Individual is climbing from one point to another in normal location (e.g. onto a platform, from top of cage, to side of cage, etc.). If this occurs with play face, then it is considered play. Includes climbing with food.
	BC	Brachiate	Animal uses arms to swing from one location to another. If this occurs with play face, then it is considered play. Includes brachiating with food.
Submissive	FL	Flee	Moving away from another at full speed. If fleeing occurs with screaming or fear grimace, will be recorded as such using modifiers.
	AV	Avoid	Moving away as another approaches or moves out of the path of another. Also includes flinching, or slight ducking of the head, or head and shoulders, or momentary retreat movements. If this occurs with a fear grimace or scream, will be recorded as such using modifiers.
	FG	Fear-grimace	Facial expression with parted lips that are pulled back, exposing the teeth which are closed. Includes bared-teeth scream, a facial expression with parted lips that are pulled back exposing the teeth accompanied by high pitched screaming.
	PT	Present	Posture may vary from extreme crouch (i.e. all four limbs are folded under such that animal is close to the ground) to smaller degrees of limb flexion. Subordinate may also turn rump towards the other without any flexion of the limbs. If this occurs with screaming, fear grimace, or pant-grunt, will be recorded as such using modifiers.

	RC	Receive contact aggression	Receiving any contact aggression behaviors.
	RA	Receive non-contact aggression	Receiving any aggression behaviors not defined as contact aggression.
	TT	Temper tantrum	Screaming loudly, may include leaping up, flinging arms above head and slapping them onto the ground, hurling self to the ground on face, hugging the cage mesh, an object or tumbling over and over, still screaming. Screaming may end in glottal cramps (gagging sounds).
Affiliative	GM	Groom mutual	Focal animal picking through hair or at skin removing debris w/ hands or mouth (not pulling) on another chimp who is concurrently grooming the focal animal.
	GG	Groom given	Focal animal picking through hair or at skin removing debris w/ hands or mouth (not pulling) on another animal.
	GR	Groom received	Another animal picking through hair or at skin removing debris w/ hands or mouth (not pulling) on the focal animal.
	EM	Embrace	Ventro-ventral contact when two individuals face each other and each puts both arms around the other or contact with both arms of an individual encompassing another from behind.
	HI	Human interaction	Non-aggressive behavior directed toward human when human is not directly interacting with chimp. Includes attempts to play with human, attempts to gain human attention (e.g. cage bangs), and attempts to give human inanimate object (e.g. twigs, wads, grass, etc.).

	PY	Play	Non-aggressive interactions involving two or more animals. Never accompanied by pilo-erection or agonism; may be accompanied by play-face and/or laughing. Includes rough-and-tumble play (fast-paced, vigorous locomotion, wrestling, hitting, pulling, chasing, biting, etc.), quiet play (slower-paced, gentle-tickling, finger and toe manipulation, etc.). May involve the use of an object.
	PS	Play-self	Individual may play quietly by itself with hands, fingers and toes, other body parts, including genitals, or an object may be handled and be the focus of play.
	OA	Other affiliative	Focal animal engaged in any affiliative behavior not described elsewhere. May include kissing, or hand to mouth contact.
Manipulate Object/Species typical	NS	Nest	Use of paper or other material to create a nest. May include smoothing material on the floor, tearing material and placing on the floor, carrying material or paper to another location.
	TU	Tool use	Includes creating and using sticks for fishing or to obtain an out-of-reach object, and using any other object or material for something other than that which it was intended.
	EO	Explore object	Sensory investigation of an object. Includes handling, touching, moving, smelling, mouthing, listening to, or tasting objects, including part of the caging structure (mesh, bolt, wood structure).
	UE	Use enrichment	Any visual, olfactory, manual, or oral use of enrichment devices, including mirrors, pipe feeders, forage boards, toys, and other enrichment objects.

Abnormal	RR	Regurgitation/ reingestion	Deliberate regurgitation and consumption of previously ingested food.
	CY	Coprophagy	Deliberate ingestion of own or other animal's feces.
	FS	Feces smear	Smearing or spreading feces on a surface with hands or mouth.
	HP	Hair pluck	Pulling out own or other's hair from the root in a quick motion, followed by manipulation of the root. Excludes self-grooming.
	IB	Idiosyncratic body manipulation	Repeated, sustained and purposeless manipulation of a specific area of own body, such as eye-poking, self-patting, or ear-covering.
	IM	Idiosyncratic movement	Repeated movement of body in purposeless manner, including neck twisting, swaying back and forth, head rolling, pacing, and rocking.
	UP	Urophagy	Ingestion of own or other animals' urine.
	OB	Other abnormal	Any abnormal behavior not described above.
Self-directed	SG	Self-groom	Picking through own hair or at skin removing debris w/ hands or mouth (not pulling). Includes visual attention directed at body part for inspection of potential grooming site.
	GS	Gentle scratch	Raking one's own hair or skin with fingernails, including mainly movements of hand and fingers. A new instance is scored after a period of 5 seconds of no scratching.
	RS	Rough scratch	Rake one's own hair or skin with fingernails, including large movements of arm. A new instance is scored after a period of 5 seconds of no scratching.

Foraging/ Feeding	FR	Foraging	Manipulating the substrate or food item while intermittently transferring items to the mouth or chewing a food item. Not stationary. Does not include chewing wads of food for long periods of time.
	ED	Eating/drinking	Consuming food items, including chow, or drinking water while remaining stationary in sitting position. Does not include chewing on wads of food for long periods of time.
Inactive	IR	Inactive rest	Animal is immobile, not engaging in activity, is generally relaxed, and is lying down, leaning, or reclining.
	IA	Inactive alert	Individual is in a sitting, prone, or quadrupedal position, is not exhibiting signs of sleepiness, and is not involved in any active behavior such as locomotion, play, eating, rocking, or grooming. Individual is alert and attention is directed toward people, chimps, inanimate object, or nearby event.
Other	PH	Pant hoot	Occurs by itself in the context of excitement. Voiced on both inhalation and exhalation and incorporates a series of "hoo" sounds which may or may not escalate to an "ahh" vocalization.
	OV	Out of view	Not visible by observer.
	OO	Other	Any behavior that is not included in this ethogram.
Sexual	CP	Copulate	Ventral surface of initiator is in contact with dorsal surface of another animal. Pelvic thrusting is exhibited. May occur with or without penetration.

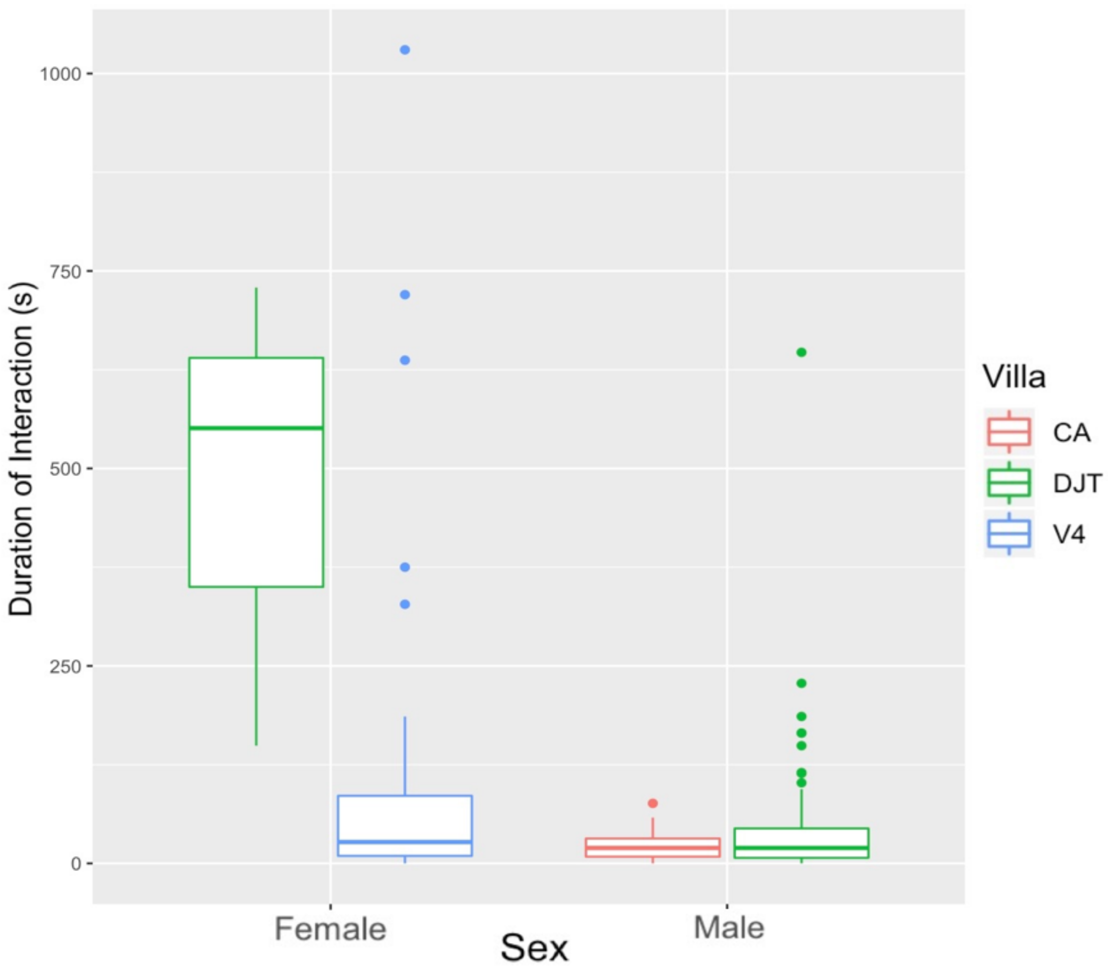
ST	Solicit	Animal (generally male) sits facing another individual with thighs rotated laterally. Male often has an erection and may exhibit thrusting motions.
MT	Mount	Ventral surface of initiator is in contact with dorsal surface of another animal with proper orientation. If thrusting occurs, will be scored as copulate. If occurs with play face, will be scored as social play.
GE	Genital explore	Visual, oral, or manual inspection of ano-genital region. May be other- or self-directed. Does not include self-grooming of the ano-genital region, manipulation of the anus to obtain feces, or masturbation.
MB	Masturbate	Using a body part, an object or a part of the cage to stimulate the subject's own genitals.

Notes. Ethogram of behaviors adapted from "Captive chimpanzee behavior as a function of space per animal and enclosure type" (Neal Webb, Hau, & Schapiro, 2018) with some definitions taken from Hopper, Freeman and Ross (2016), and AZA Ape Taxon Advisory Group & AZA Animal Welfare Committee (2010).

B3

Total Box Interactions and Stress-Related Behaviors

ID	Explorations	Uses	Duration	SRB	Aggressions	SDB
BC	1	1	551	1	1	0
CS	1	0	149	1	1	0
EM	1	1	729	1	1	0
BI	11	1	199	1	1	0
DR	3	0	113	0	0	0
HC	10	1	563	1	0	1
JC	2	0	64	2	2	0
KD	8	1	261	3	1	2
KU	10	3	814	3	2	1
LE	21	2	1349	1	0	1
OC	7	1	152	2	2	0
RY	5	0	57	0	0	0
AU	2	0	33	0	0	0
JH	4	0	114	0	0	0
MR	3	0	56	0	0	0
PT	3	0	108	0	0	0
QU	3	2	107	0	0	0
TZ	1	0	28	0	0	0
AL	10	0	478	1	1	0
AY	12	3	834	1	0	1
BB	6	0	354	0	0	0
HR	9	1	291	1	1	0
LB	6	0	328	1	1	0
LO	8	3	612	0	0	0
NO	20	16	3522	27	26	1
SR	15	0	357	1	1	0
SY	3	0	11	0	0	0



B4. Mean durations of box use by sex and villa.