Spring 5-11-2018

The Conundrum of Causal Reasoning in Elephants

Beri Brown

Follow this and additional works at: https://aquila.usm.edu/masters_theses
Part of the Cognition and Perception Commons, and the Comparative Psychology Commons

Recommended Citation
https://aquila.usm.edu/masters_theses/340

This Masters Thesis is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Master's Theses by an authorized administrator of The Aquila Digital Community. For more information, please contact Joshua.Cromwell@usm.edu.
THE CONUNDRUM OF CAUSAL REASONING IN ELEPHANTS

by

Beri Brown

A Thesis
Submitted to the Graduate School,
the College of Education and Psychology,
and the Department of Psychology
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Master of Arts

May 2018
THE CONUNDRUM OF CAUSAL REASONING IN ELEPHANTS

by Beri Brown

May 2018

Approved by:

________________
Dr. Alen Hajnal, Committee Chair
Associate Professor, Psychology

________________
Dr. Richard Mohn, Committee Member
Associate Professor, Educational Research and Administration

________________
Dr. Heather Hill, Committee Member
Associate Research Professor, Psychology

________________
Dr. D. Joe Olmi
Chair, Department of Psychology

________________
Dr. Karen S. Coats
Dean of the Graduate School
ABSTRACT

THE CONUNDRUM OF CAUSAL REASONING IN ELEPHANTS

by Beri Brown

May 2018

Causal reasoning is marked by the ability to mentally reconstruct the missing part of a sequence in order to reproduce an outcome. While research on causal reasoning has been done with children, the results of the studies have been inconsistent. A standardized paradigm for comparative causal reasoning studies does not exist. Nissani (2006) investigated causal reasoning in a tool-use task with elephants and concluded that elephants were not capable of causal reasoning. The current study, a modified replication, yielded results that were not congruent with Nissani’s (2006) manuscript. Additionally, it was very unlikely that the Nissani (2006) study truly looked at causal reasoning or tool-use, and instead assessed a response acquired through associative learning. Based on the results of the current study, it appears that elephants are capable of a level of causal reasoning, although more research is necessary.
ACKNOWLEDGMENTS

It really does take a village and I am so grateful to have the most incredible village. To my committee: thank you! Dr. Richard Mohn, for your support and advice, Dr. Alen Hajnal for your patience, encouragement, and very appreciated dissemination of information, and to Dr. Heather Hill: my academic saving grace! You have been my biggest cheerleader, sounding board, and all around rock star throughout this process. I will always be beyond thankful for everything you have done for me, and for understanding all my quirky Beri-isms.

Nick Way, I literally could not have done this without you! Thank you for being so amazing with the elephants, understanding of the scientific process, sweating in Salinas, putting me in touch with Charlie, your patience with even the long trials, and for all of your support and encouragement in making all of this happen! To Kara, Andrew, Brett, and Robert, and Patrick: thank you for working the elephants, going on an adventure to Salinas, and for making the jazziest food rewards for the elephants. Dianne Cameron, thank you for all of your guidance and help as well. The entire SFDK family has been so phenomenal: jazz pec-ies to all of you! And to Charlie Sammut, thank you so much for letting us borrow your elephants for the day!

Dr. Pepper Hanna, thank you for being my MMBCL rock and for being willing to answer all of my questions, regardless of how ridiculous they are, and for cheering me on throughout our entire graduate journey. I would not have made it without you! To my brother, Dr. Dimitri Brown, for graciously doing my reliability coding, (and my sister Ariana Brown for also offering to do coding)!
And to my parents who are the most incredible, patient, and loving support system. Thank you for always supporting me and encouraging me to pursue (all of) my dreams! I absolutely could not have done with without you! Mom, you’ve been my rock through everything, and I’m endlessly thankful! And to my little dog, Rora, you’ve literally been by my side every step of the way. Thank you for all of the cuddles and unconditional love!
DEDICATION

I would like to dedicate this manuscript to Dr. Stan Kuczaj. Stan, thank you for taking a chance on me, believing in me from the start, and always supporting me for being myself. While I wish that you were still here so we could talk about Texas, food, football, and marine mammal cognition, I am so honored and thankful for the time I got to spend under your guidance as a member of the MMBCL. From the trip to Honduras and The Bahamas to Boat Trips and lab meetings at The Keg… even our lab meetings and your lectures about developmental psychology. Thank you for all of the laughs, the fist bumps, the adventures, the knowledge, and for being such a huge part of shaping who I am as a scientist! You will always be missed, but also remembered for the incredible person you were!
## TABLE OF CONTENTS

ABSTRACT .......................................................................................................................... ii

ACKNOWLEDGMENTS ................................................................................................. iii

DEDICATION .................................................................................................................. v

LIST OF TABLES ........................................................................................................... viii

LIST OF ILLUSTRATIONS ............................................................................................. ix

LIST OF ABBREVIATIONS ............................................................................................ x

CHAPTER I - INTRODUCTION .......................................................................................  1

What is Causal Reasoning? ..........................................................................................  1

Current Methods Used to Assess Causal Reasoning in Nonverbal Individuals ..........  2

Nonverbal Humans. .......................................................................................................  2

Nonverbal Animals. .......................................................................................................  5

Causal Reasoning and Tool Use ..................................................................................  6

Human Tests. ................................................................................................................  7

Non-Human Animal Tests. ..........................................................................................  8

Trap-tube test. .................................................................................................................  8

Avian Species .................................................................................................................  9

Non-Human Primates ...................................................................................................... 10

Comparative Studies with Primates and Humans. ......................................................  12

Tool Use and Causal Reasoning in Dolphins. ..............................................................  13
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool Use in Elephants.</td>
<td>14</td>
</tr>
<tr>
<td>Causal Reasoning and Tool Use in Elephants.</td>
<td>15</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>16</td>
</tr>
<tr>
<td>Limitations of the Original Study.</td>
<td>16</td>
</tr>
<tr>
<td>Modifications to Original Study.</td>
<td>17</td>
</tr>
<tr>
<td>CHAPTER II - METHODS</td>
<td>20</td>
</tr>
<tr>
<td>Subjects of Archived Data</td>
<td>20</td>
</tr>
<tr>
<td>Experimental Testing Apparatus.</td>
<td>20</td>
</tr>
<tr>
<td>Experimental Procedure</td>
<td>22</td>
</tr>
<tr>
<td>Experimental trials</td>
<td>22</td>
</tr>
<tr>
<td>Coding of trials</td>
<td>24</td>
</tr>
<tr>
<td>Statistical Analyses</td>
<td>25</td>
</tr>
<tr>
<td>Intercoder Reliability</td>
<td>26</td>
</tr>
<tr>
<td>CHAPTER III - RESULTS</td>
<td>28</td>
</tr>
<tr>
<td>CHAPTER IV – DISCUSSION</td>
<td>30</td>
</tr>
<tr>
<td>APPENDIX A – IRB Approval Letter</td>
<td>43</td>
</tr>
<tr>
<td>APPENDIX B – SFDK Approval Letter</td>
<td>44</td>
</tr>
<tr>
<td>APPENDIX C - MZ Approval Letter</td>
<td>45</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>46</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1 Comparison of Nissani (2006) Study and Current Study Methodologies ........ 19
Table 2 Comparison of Overall Results................................................................. 38
LIST OF ILLUSTRATIONS

Figure 1. Front trial at SFDK. ................................................................. 27

Figure 2. Top trial at MZ................................................................. 27
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFDK</td>
<td>Six Flags Discovery Kingdom</td>
</tr>
<tr>
<td>MZ</td>
<td>Monterey Zoo</td>
</tr>
</tbody>
</table>
CHAPTER I - INTRODUCTION

What is Causal Reasoning?

Causal reasoning is marked by the ability to mentally reconstruct the missing part of a sequence in order to reproduce an outcome. Premack and Premack (1994) refer to this as “Sherlock Holmes” style reasoning in which the “detective” has to put together and recreate missing “puzzle pieces” to determine how and why an outcome occurs, using only current or previous observations and experiences. Several components are needed to determine causal reasoning. The first component is an association between the antecedent and consequent events formed when events are paired together because of their spatio-temporal proximity to one another rather than due to a cause and effect relationship (Visalberghi, & Tomasello, 1998). The second component extends this simple association into the development of an elucidatory hypothesis about the potential cause of the experienced or desired effect. The final component requires the formulation of potential mediating forces that enabled the cause and effect outcome. In short, causal reasoning necessitates comprehension, prediction, postdiction, and, in most experimental settings, ultimately production (Visalberghi, & Tomasello, 1998). Simply put, “causal knowledge allows us to predict outcomes and underlies our ability to control events in the world” (Blaisdell, Sawa, Leising, & Waldman, 2006, p. 1020).

The differentiation between associative learning and causal learning, especially in comparative studies, can be difficult to tease apart. While associative learning pairs objects based on their temporal relationship, causal reasoning requires the additional understanding of how and why objects are paired (i.e., explaining their causal relationship and the underlying correlation; Taylor, Hunty, Median, & Gray., 2009;
Zuberbu, 2000). Premack and Premack (1994), noted in their comparative studies with chimpanzees and children that the key to causal reasoning was reconstructing missing information to identify the reason or motivation behind the relationship between two objects. Associative learning does not consider the reason or motivation behind the relationship between two objects.

Current Methods Used to Assess Causal Reasoning in Nonverbal Individuals

Nonverbal Humans.

Demonstrating a capacity for causal reasoning is challenging, especially in nonverbal humans and animals. Most of the research on causal reasoning that has been conducted with children used slightly older children who were capable of articulating their thoughts and logic, or at least a few words to the experimenter (Premack & Premack, 1994). Infants, however, take the language component out of the equation. Consequently, the methodologies used with infants are more comparable to that which are used with nonverbal animals. On the other hand, the lack of muscular strength, dexterity, locomotion, and overall muscular control of human infants presents additional challenges that are not present in juvenile or adult animals. These limitations have resulted in predominately passive observational roles when researching human infants. Research with gaze trackers has been one of the more predominant means of collecting data to better understand causal reasoning in human infants.

Due to their limited mobility and muscular abilities, one of the most common means of looking at causal reasoning in infants is by tracking their gaze or attention. Causal reasoning studies with infants typically involve children looking at a screen that presents either animated schematic events (Leslie, 1982; Leslie & Keeble, 1987; Oakes &
Cohen, 1990; Schlottmann & Surian, 1999) or schematic events performed with puppets and toys (Muentener & Carey, 2010; Saxe et al., 2007). These studies suggest that infants possess a level of causal understanding that appears to develop after six or more months of life (Leslie, 1982; Leslie & Keeble, 1987; Muentener & Carey, 2010; Oakes & Cohen, 1990; Saxe et al., 2007; Schlottmann & Surian, 1999). The results of these experiments preliminarily support the presence of causal reasoning in humans as young as 27 weeks old (Leslie & Keeble, 1987), but more research is still needed to fully understand the true origin of humans’ capacity for causal reasoning (Saxe & Carey, 2006).

Unlike paradigms with an infant as a passive observer, one paradigm allowed young children to take a more active role, similar to that used with animals, called the “blicket detector” (Gopnik & Glymour, 2002; Gopnik & Sobel, 2000; Gopnik, Sobel, Schulz, & Glymour, 2001; Gopnik et al., 2004; Sobel & Kirkham, 2006, Sobel & Krikham, 2007). The blicket detector was a machine that lit up and played music when a predetermined object/blicket was placed in front of it. Participants were individually shown two objects (A and B). One object, the blicket, would cause activation of the detector, while the other would not. When first shown the paradigm, the child would watch the experimenter place both objects in front of the detector together (activating the detector). Next, only one of the objects (A) would be placed in front of the detector and that object would not cause a reaction, which should allow the child to infer that the other object (B) was the blicket and caused the reaction when they were presented together. Consequently, when asked to light up the detector, the child was expected be able to select only the blicket that lit up the detector. This procedure was then repeated using
other blickets and objects that varied in shape and color. The blicket that caused the reaction was randomized across trials. The original experiment by Gopnik and Sobel, (2000), demonstrated that two-, three-, and four-year-olds were capable of making these causal inferences. Sobel and Kirkham (2006) set out to determine if younger babies had comparable causal abilities, and using a similar paradigm with 19- and 24 – month-olds. Ultimately, they found that both 19- and 24 – month-old babies were capable of this initial causal understanding, but that the 24-month-olds had a more robust understanding of the causal relationship when the paradigm became more complex (Sobel & Kirkham, 2006). Wanting to extend this paradigm to even younger babies, Sobel and Kirkham (2004) used a similar paradigm with 8-month-olds, in which anticipatory eye movements were measured. These infants were presented with a series of four sequenced events. When the trained sequence was changed, the dependent variable became the amount of time the infant spent looking at certain frames, which revealed the infant’s anticipated sequential pattern. These results indicated the existence of causal understanding in infants similar to that found in the blicket studies with 24-month-olds and the original Gopnik and Sobel (2000) study with preschoolers. Sobel and Kirkham (2007) also used the same visual sequencing paradigm used with eight-month-olds to test five-month-olds, and had similar results.

These findings on causal knowledge, using either the blicket detector or the visual tracking paradigm, suggest that learners can differentiate two events based on a direct relationship between them and can also determine that events are independent based on the value of other events (Gopnik & Glymour, 2002; Gopnik & Sobel, 2000; Gopnik et al., 2001; Gopnik et al., 2004; Sobel & Kirkham, 2006, Sobel & Krikham, 2007). These
findings have been tested with Bayes nets, which are based on the Markov assumption derived from a graphical modeling framework. Causal Bayes nets, based on probabilities, are theoretical probabilistic graphical models of causal relationships that impact causal learning (Hagmayer, 2016; Gopnick et al., 2004). These graphical models have been used to statistically predict and demonstrate causal knowledge in young children (Gopnik & Glymour, 2002; Gopnik & Sobel, 2000; Gopnik et al., 2001; Gopnik et al., 2004; Sobel & Kirkham, 2006; Sobel & Krikham, 2007).

Nonverbal Animals.

One of the more compelling arguments for a species, other than humans, having the capacity for causal reasoning comes from research with rats (*Rattus norvegicus*) by Blaisdell and his colleagues (2006). Similar to that which was done with infants and toddlers (Gopnik & Glymour, 2002; Gopnik & Sobel, 2000; Gopnik et al., 2001; Gopnik et al., 2004; Sobel & Kirkham, 2006, Sobel & Krikham, 2007), Blaisdell et al. (2006), also used causal Bayes nets to represent the causal relationships of variables in their experiments with rats. More specifically, Blaisdell et al. (2006) looked at Pavlovian conditioning in rats and the impact of seeing (observations) and doing (interventions) on the rat’s causal model. Initially, the rats observed a light which proceeded a tone and the simultaneous production of a sucrose solution. Once this association was made, the rat interacted with the paradigm and learned that pressing a lever controlled the production of the tone. The causal Bayes net models suggested that this would cause the dissociation between the tone and the food, while the association between the light and the food should remain intact. Associative behavioral models, however, suggested that the rat’s behavior would not differ based on whether the rat was an observer or actively
intervening. Ultimately, the predictions of the Bayes nets were supported. The authors suggest that this paradigm may have been successful in part because it was not dependent on the rats having a vast physical knowledge that they might not possess.

Rats, although easy to study in large quantities in laboratory settings, are not usually regarded as highly intelligent or cognitive animals. With regards to brain weight, encephalization quotient (EQ), and number of cortical neurons, which are all hypothesized to correlate to intelligence, rats and humans are at opposite ends of the spectrum (Hart, Hart, & Pinter-Wollman, 2008; Roth & Dicke, 2005). Closest to humans, with regards to EQ are bottlenose dolphins, followed by non-human primates, while the heaviest brains are found in elephants and whales. Elephants have the highest number of cortical neurons, second to humans (Roth & Dicke, 2005). Consequently, while the rat research is intriguing, it is probable that cognitive research with species regarded as having higher levels of intelligence and cognition would be more relevant, especially in comparative studies.

Causal Reasoning and Tool Use

According to Piaget, 1954, infants begin learning about their world through trial and error. Initially unable to predict the outcome of behavior, infants start to develop causal knowledge of their own body and how it interacts with the world around them through their self-exploration of their body’s actions and subsequent responses. As human infants mature, the corresponding change in their cognitive abilities helps them to remember these causal connections. If an event produced a positive result, infants try to reproduce that result repeatedly. Early on, everything is based on physical interactions. As the infant’s motor control increases, he or she can start to intentionally interact with
his or her world based on the causal knowledge he or she has built through observation and interaction. Consequently, without the causal knowledge gained as infants, children would not know how to interact with their world and consequently how to use their body or tools. Accordingly, causal reasoning can be thought of as a precursor of tool-use, since without it, tool-use could not successfully take place (Piaget, 1954; Schlesinger & Langer, 1999). Tool-use, defined by Shumaker et al. (2011), is:

The external employment of an unattached or manipulable attached environmental object to alter more effectively the form, position, or condition of another object, another organism, or the user itself, when the user holds and directly manipulates the tool during or prior to use and is responsible for the proper and effective orientation of the tool. (p. 5)

This definition requires that an unattached object, or attached manipulable object be used to influence another object, requiring the user to have an understanding of how the tool will have an effect on the object in order to reproduce this manipulation.

*Human Tests.*

To examine the development of tool use in children, Schlesinger and Langer (1999) presented eight- and 12-month-olds with two types of tool use tasks. Each task had one condition that was possible to complete and another that was not possible. The first task involved retrieving a toy that was placed on top of or to the side of a cloth. Successful completion indicated that the child either understood that the cloth had to be supporting the toy and then it could be used as a tool to retrieve the toy. Another task utilized a hook that was either surrounding the toy or to the side of the toy. This task
required the child to recognize that the hook had to be around the you in order to retrieve it. As was hypothesized, both ages correctly completed the cloth task, but only the 12-month-olds had the causal knowledge to correctly use the hook tool. Schlesinger and Langer (1999) also looked to see if the causal knowledge of the children changed based on whether they were observing the trials or an active agent in the trials. This study revealed that eight-month-olds did not visually predict which cloth trials would be successful, although 12-month-olds did have this capacity. Neither 12- nor eight-month-olds visually predicted which hook positions would be successful. This pattern suggests that whether the participant is an observer or an active agent in a trial can have an impact on the participant’s causal understanding. When actively interacting with the paradigm, the children were better able to predict the outcome of trials than when they merely observed.

**Non-Human Animal Tests.**

*Trap-tube test.* There does not appear to be a standard, universally accepted, paradigm for causal reasoning when testing non-human animals. Nonetheless, the tube-trap test has been tested across multiple species including woodpecker finches *(Camarhynchus pallidus)*, rooks *(Corvus frugileus)*, crows *(Corvus moneduloides)*, and non-human primates *(Premack & Premack, 1994; Seed, Tebbich, Emery, & Clayton, 2006; Silva, Page, & Silva, 2005; Taylor, Hunt, Medina, & Gray, 2009; Tebbich & Bshary, 2004; Visalberghi & Limongelli, 1994; Visalberghi & Tomasello, 1998; Zuberbuhler, 2000).*

There are several variations of the trap-tube test, but the premise of the test involves the placement of food in a horizontal tube, which has a hole in the bottom that
leads to a recessed container or trap. Participants must use a tool (usually a stick) to push or pull the food out of the tube in the correct direction, avoiding having the food fall in the trap. In addition to potentially requiring the subject to have a basic knowledge of the causal relationships involved in the trap-tube test, this paradigm requires that the subject has at least a basic understanding of tool use. Consequently, the subject must be able to not only use a tool to complete the task, but the subject must also be able to causally understand what impact his or her action with the tool will have on his or her ability to retrieve the reward. This requires a causal understanding of the trap and a mental representation of how to both retrieve the reward and avoid the trap.

*Avian Species.* Avian species have been the subjects for several trap-tube studies. Much like rats, birds are easy to house in laboratories. Furthermore, their capacity for cognitive tasks, such as tool use, makes them a relevant candidate for research on causal reasoning using tools.

Tebbich and Bshary (2004) used the trap-tube test paradigm with woodpecker finches (2 juveniles, 4 adults). Only one woodpecker obtained the food reward well above chance, while the other five were significantly below chance. When given a task where the tool had to be modified to be used, four of the five modified the tool and retrieved the food on the first trial, but did not show a decrease in errors as the trials progressed (indicative of learning/causal understanding). As a result, no conclusive results can be drawn from this study.

Using a similar paradigm, Seed et al. (2006), looked at the same ability in eight rooks, with a trap-tube test that had two traps. Initially, both traps were covered so that the rook had to learn to go over the plugged trap. Seven of the eight rooks were able to
complete this task. In the second setup, one trap was covered at the bottom so that the food could fall into it but could not be retrieved, and the other was uncovered so that the food could fall out of it and be eaten. All seven of the rooks who completed the first variation of the test were able to complete this one as well. Two additional variations were also tested, but only one rook succeeded in all four paradigms. Additional research is still needed to be able to draw conclusive results about causal reasoning in rooks.

Taylor et al. (2009), used both the trap-tube test and the trap table test with crows. These tasks were visually distinct, but causally equivalent, and therefore controlled for the possibility of their results being due to associative learning. The researchers created tasks that differed in shape, color, and material. Six wild Caldonian crows, five adults and one sub-adult, were tested. The initial trap-tube test was the two-trap test, similar to the task used for the rooks (Seed et al., 2006). Crows were tested with several variations of the trap-tube test and three of the six crows were successful. Next, the crows were given a trap table, requiring them to choose between using a tool to pull food along a continuous surface and pulling the meat from behind a rectangular trap, into which the meat would fall and no longer be retrievable. All three crows solved the trap-table test on their first trial. In order to ensure that the crows’ success was not the result of an innate instinct to avoid holes, the three crows who failed the first trap-tube test were tested on the table-trap test. They were not successful. The results were variable across subjects, indicating that further exploration is still needed before robust conclusions can be drawn.

*Non-Human Primates.* Although multiple species of primates have been tested, there appear to be no consistent trends in their abilities. As a result, the trap-tube test does not appear to be a reliable measure of causal reasoning. Nonetheless, the trap-tube
test has been the most frequently used paradigm for investigating causal reasoning in primates and in comparative studies with humans and primates.

One of the first studies to look at non-human primate understanding of causal reasoning using the tube-trap paradigm was conducted by Visalberghi and Trinca (1989), tested four tufted capuchins (Cebus apella), but modified the paradigm to use a regular pvc pipe without a trap. The capuchins had to determine the proper tool to use to retrieve (via pushing or pulling the food) the reward. Three of the four used the sticks to push the food reward out without any additional direction, but it soon became apparent that their understanding of the paradigm was very limited and a disparity between the ability to solve the task and an understanding of the task was revealed, rendering the outcome inconclusive.

Visgalberghi and Limongelli (1993) added a trap in the center of the tube, to determine if the capuchins could understand the effect that their actions had on the movement of the food reward. They tested four capuchins, and only the youngest appeared to understand the paradigm. However, when the tube was inverted, the capuchin still acted as if the reward could fall in the trap. This response suggested that the capuchins did not have a full causal understanding of the paradigm, even if they could sometimes correctly solve the task. When a similar paradigm was conducted with five chimpanzees (Pan troglodytes), only two of the chimpanzees performed above chance (Limonigelli, Boysen, & Visalberghi, 2005).

Povinelli (2000) tested seven chimpanzees with an inverted trap tube test, in order to determine if the chimpanzees fully understood the paradigm and subsequently causal reasoning, or the animals were using some other rule. Povinelli (2000) discovered that
the chimpanzees pushed the reward away from the trap, even when the tube was inverted and the food could not fall into the trap. This standardized behavior suggested that the chimpanzees did not actually understand the paradigm, but rather, they had found a rule that consistently yielded positive results.

Five orangutans (Pongo pygmaeus), two chimpanzees, two bonobos (Pan paniscus), and one gorilla (Gorilla beringei graueri) were also given the trap-tube test that required that a choice be made between pushing and pulling the food reward (Mulcahy & Call, 2006). With the exception of one bonobo, all of the primates preferred raking the reward toward themselves over pushing the reward away. Of the 10 subjects, only one chimpanzee and two orangutans were able to solve the right-side-up trap tube paradigm above chance. When the tube was inverted, they no longer avoided the trap. Like many trap-tube tests, this study yielded mixed results.

*Comparative Studies with Primates and Humans.*

The inability of primates to successfully complete the trap-tube test surprised many researchers, perceiving the task as seemingly simple. In an attempt to illuminate why it was so challenging, Silva et al. (2005) did a comparative study with chimpanzees and adults. Ten undergraduate students were first tested with three different tube apparatuses: one with no traps, one with a single trap, and one with two traps (Silva et al., 2005). These experiments were also performed with opaque (instead of transparent) tubes, and a mark was placed on the outside of the tube to indicate where the reward had been placed. Some problems had no solution (e.g., reward placed between the traps), while others had multiple potential solutions. The first experiment revealed that humans preferred to insert the stick into the end that would cause the reward to 1) not fall in the
trap and 2) travel the least distance. In experiment two, 24 undergraduates were shown drawings of each of the 15 different combinations used in experiment one (reward placement and tube type were randomized). They were then given an answer sheet and asked to select left or right (indicating the side from which the stick should be inserted). Two significant differences emerged. A bias toward inserting the stick from the left side appeared, both when there was no trap, or when it was inverted and did not affect the outcome, and when the reward was equidistant from the ends. In their third experiment, they looked to see if people were more likely to pull a rake on the side of a table that did not have a hole to retrieve a reward, or if they would pull a rake on the side of a table that had a hole, but in an ineffective place (so the reward would be received in both scenarios). Nineteen undergraduates were tested with 15, for one paradigm, and 16, for another paradigm, preferentially choosing the side of the table with no hole. Primarily, humans responded by trying to avoid the possibility of the food falling in the hole. Some of these behavioral biases could be construed as a lack of understanding of the paradigm, but fortunately humans were able to explain the reasoning behind their decisions. This qualitative approach illuminated the possibility that behavioral biases might overshadow and confound causal knowledge. The dearth of conclusive results, even for adults who could articulate a level of understanding of the paradigm, but still had behavioral patterns inconsistent with those expected, supports that there does not seem to be a valid and reliable paradigm in place yet for the study of causal reasoning.

*Tool Use and Causal Reasoning in Dolphins.*

Dolphins, with Eqs most similar to humans, have also been used in a paradigm similar to the trap-tube test, but adapted for flippers and being underwater. Kuczaj and
colleagues (2009), highlighted the possibility of causal reasoning in Atlantic bottlenose dolphins (*Tursiops truncatus*). Using a Multiple Weight Test with two dolphins at Disney’s Living Seas at EPCOT, the dolphins had to learn to drop four weights in to an apparatus to release a fish that could be eaten. To successfully complete this task, the dolphins had to plan their actions with weights, which are likely predicated on an understanding of causal reasoning and planning, as proposed by Kuczaj, Gory, and Xitco (2009). The dolphins successfully completed the task. However, since causal reasoning was not the main focus of the study, further research is needed to conclusively say that dolphins are capable of causal reasoning. Nonetheless, dolphins’ success with the Multiple Weight Test could be argued as support for causal reasoning using tools.

**Tool Use in Elephants.**

Elephants, with the heaviest brains, second only to some species of whales, are a strong candidate for cognitive research (Hart, et al., 2008). It has been reported that elephants use tools with the highest frequency and diversity of non-primate humans (Beck, 1980; Chevalier-Skolnikoff & Liska, 1993). Captive elephants have shown 21 types of tool use, falling into three categories: body care, aggression and ambiguous, while wild elephants showed nine types of tool use that fell into four contexts: body care, aggression, feeding another, and ambiguous behavior toward a dead conspecific (Chevalier-Skolnikoff & Liska, 1993). Regardless of setting, 80% of elephant tool use occurs during body care (Chevalier-Skolnikoff & Liska, 1993). Elephants constantly manipulate and explore things with their trunk, a prehensile organ, with a finger-like protrusion at the end of it that is often used to manipulate objects and acquire food. However, elephants use branches as tools for scratching their body, throwing at other
animals or people, and for fly switching (Chevalier-Skolnikoff & Liska, 1993; Hart & Hart, 1994; Hart, Hart, McCoy, & Sarath, 2001). Consequently, it can be inferred that tool-use is not only something elephants are capable of, but also part of their natural repertoire such as assisting in essential functions like body care and foraging.

Causal Reasoning and Tool Use in Elephants.

To date, only one study has been conducted to experimentally examine causal reasoning in elephants with a tool-use task (Nissani, 2006). Four female Asian elephants (Elephas maximus) were initially trained to throw the lid off a bucket to obtain the food reward inside the bucket. Once the elephants were trained on this task, experimental trials began in which the lid was placed on the ground to the right or the left side of the bucket. Trials were then randomized between having the lid on top, to the right, and to the left of the bucket. The elephants’ behavior was analyzed to determine if elephants still touched with the lid on the side of the bucket, or if the elephants retrieved the food without interacting with the lid. The elephants were only given the opportunity to obtain a food reward if the lid was not thrown. In the first five side-placement trials, all four elephants interacted with the lid. By the last six experimental trials, only the youngest elephant retrieved the food without touching the lid, while the other elephants still tossed the lid before retrieving the food more often than not. The number of trials varied between elephants. This experiment was repeated with 11 different elephants (ages 12-17) and a slightly modified paradigm (with regards to the behavior of the experimenter). The first five side placement trials were the focus of the analysis. Out of these first 55 trials (5 per elephant), the lid was only ignored in three trials. Nissani (2006) argued that since elephants continued to throw the lid when it was not an obstacle to the food, the
elephants were not capable of causal reasoning because they did not appear to understand the causal relationship between the food and the lid. If the elephants had understood this relationship, they would have ignored the lid as it was not necessary to complete the task.

**Purpose of the Study**

The purpose of the study was to use archival data to reassess Nissani’s (2006) manuscript about causal reasoning in elephants. To date, Nissani’s study is the only available data on causal reasoning in elephants. Unfortunately, this study had many methodological flaws that likely confounded the interpretation of the findings.

Although acknowledged by the author, one possible confound was related to the previously trained and continually reinforced behavior of the lid throwing by the elephants. The elephants were trained to throw the lid off the bucket in their original trials and consequently, likely carried that behavior into subsequent trials. This action does not necessarily mean that the elephants did not understand that there was no food under the lid when it was not on the bucket. Other authors, when referencing the Nissani (2006) have also pointed out the potential confound of the heavy training the elephants received (Plotnik, de Waal, & Reiss, 2010). Furthermore, it is mentioned that concurrent experiments were performed between the lid-throwing sessions, using the same materials, and a desired response of lid throwing. These additional studies may have interfered with the elephants’ ability to produce the desired behavior clearly, thus producing another potential confound.

*Limitations of the Original Study.*

The original elephant lid throwing study was based on learning/conditioning explanations and not causal reasoning explanations (Nissani, 2006). It almost appears
that the study examined the extinction curve of a trained behavior rather than the causal reasoning explanation as implicated by the title and conclusions presented. In fact, when reviewing the Nisanni (2006) study, other authors put the term causal reasoning in quotations, suggesting that it is not really the most parsimonious interpretation of the data (Hart et al., 2008, p. 88, Plotnik et al, 2010, p. 182). Furthermore, to achieve an extinction curve of a trained behavior, the behavior can no longer be reinforced, which was a major confound in the Nissani (2006) study. The lid throwing behavior was both required and reinforced when the lid was on the bucket, as well as in the aforementioned concurrent research experiments.

Modifications to Original Study.

In an attempt to eliminate pre-training confounds, the elephants in the current study were not trained to remove the lid. This naiveté allowed the elephants to have an unbiased interaction with the trashcan, which minimized the influence of any previous training or interactions with the trashcan or the lid. Additionally, the keepers were always standing behind the elephant’s line of vision while the elephant interacted with the trashcan, so that they would not inadvertently cue the elephant’s behavior or tarnish the elephant’s capacity for an uninfluenced interaction with the trashcan and the lid, which was another concern with the Nissani (2006) study. The elephants in the current study also did not engage in any other concurrent experimental trials. All trials were completed in one session on one day.

To this end, concerns arise from whether the Nissani (2006) study actually addressed causal reasoning or tool use. Premack and Premack (1994) state that causal reasoning must require reasoning to understand the relationship between cause and effect.
This is not really applicable in the lid paradigm. If the cause is perceived as throwing the lid, and the effect is acquiring the reward, then 2/3 of the trials should be performed without the cause. This paradigm does not fit the definition of causal reasoning and instead is more in alignment with the definition of associative learning, where a relationship is made based on the temporal proximity of events (Taylor et al., 2009; Zuberbu, 2000).

Likewise, whether or not there is tool use in the original study must also be taken into consideration. According to the Shumaker et al. (2011), definition of tool use, there is no tool involved in the Nissani (2006) study. The only two moveable components in the Nissani (2006) study are the lid and the elephant’s trunk. The lid is a potential obstacle, and is not used in two thirds of the trials. Consequently, it should not be regarded as a tool. Similarly, the elephant’s trunk is part of the elephant’s body, so it is not a tool either. As a result, it would appear that there was no tool use in the Nissani (2006) study and tool use should be removed from the discussion and interpretation of the original experiments.

Since the Nissani (2006) study was the only study on causal reasoning (with a tool-use task) on elephants that has been published, it is important to point out these shortcomings and investigate whether methodological errors and misguided interpretations can account for the behavioral patterns of the elephants (summarized in Table 1). It was hypothesized that an in-depth evaluation of a more controlled experiment would yield behavioral patterns that suggest the elephants’ behavior was the result of their initial interaction with the lid, instead of a lack of causal reasoning.
Consequently, the proposed study was performed to examine if the original placement of the lid impacted an elephant’s behavior on subsequent trials, using an untrained task.

Table 1

*Comparison of Nissani (2006) Study and Current Study Methodologies*

<table>
<thead>
<tr>
<th>Did training/pre-training occur?</th>
<th>Nissani Study</th>
<th>Current Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes, as well as exposure to the paradigm for other experiments.</td>
<td>No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Were all trials performed in one session?</th>
<th>Nissani Study</th>
<th>Current Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. Trials happened in multiple sessions over multiple days, while also intermittently participating in other experiments that involved the same materials.</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Nissani Study</th>
<th>Current Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experiment 1: n=4 (4 female Asian elephants)</td>
<td>SFDK: n=4 (2 female Asian, 2 female African elephants)</td>
</tr>
<tr>
<td></td>
<td>Experiment 2: n = 11 (4 male, 7 female Asian elephants)</td>
<td>MZ: n = 5 (4 female African, 1 male African elephants)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experimental setup</th>
<th>Nissani Study</th>
<th>Current Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Used elements from other concurrent experimental paradigms. Bucket with lids of varying colors and boxes of varying sizes.</td>
<td>Same paradigm at both facilities SFDK – curved lid MZ- curved lid</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Nissani Study</th>
<th>Current Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Varied from trial to trial and day to day. Author mentions potential bias that was accounted for in second experiment. Research was positioned at the bucket or box, facing toward the elephant, to manipulate the placement of the food reward</td>
<td>Consistent throughout experiment. Researcher was always behind the elephant’s eye line during trial, unless guiding the elephant’s attention</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experimental controls</th>
<th>Nissani Study</th>
<th>Current Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Limited. Elephants varied greatly in the timing and number of trials, experimental paradigm exposure, and experimental apparatus</td>
<td>Each elephant assigned one of two treatments (top first or front first) and participated in only 15 trials</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Were the elephants trained free contact?</th>
<th>Nissani Study</th>
<th>Current Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
CHAPTER II - METHODS

An archived set of video recordings of the experimental trials was coded for the proposed study. The experiment was conducted in 2011, with the permission of the animal research committee of Six Flags Discovery Kingdom (SFDK), in Vallejo, CA and Monterey Zoo (MZ) in Salinas, CA. A total of 135 trials are available, with 15 trials were conducted per animal.

Subjects of Archived Data.

The subjects at SFDK included two female Asian elephants (*Elephas maximus*): Liz and Bertie Mae, ages 31 and 47, and two female African elephants (*Loxodonta africana*): Valerie and Tava, ages 29 and 33. All elephants were all housed together. All of the elephants were trained to perform in shows and give elephant rides, in addition to husbandry behaviors. The elephants had previously participated in research experiments, although none of them had been trained to manipulate a trashcan lid.

The subjects at MZ included one male African elephant: Butch, age 30, and four female African elephants: Christie, Malika, Paula, and Buffy, ages 40, 24, 40, and 29. These elephants were housed together, but did not perform in shows or give elephant rides. They were, however, trained to perform basic husbandry behaviors. They had not previously had any training with trashcans or trashcan lids.

Experimental Testing Apparatus.

The testing apparatus was comprised of a 32-gallon Rubbermaid Roughneck trashcan and its corresponding lid. At SFDK this lid was curved, while at MZ the lid was flat. At SFDK, the trashcan was chained to the metal pole of an outdoor pavilion behind
the elephants’ habitat, in a “backstage” area, using a metal chain fastened with a carabineer. At MZ, the trashcan was chained to the corner metal posts of an empty outdoor elephant corral (see Figure 1 for SFDK and Figure 2 for MZ). Having the trashcan chained to the pole allowed the elephants to manipulate the can freely, without any concern of their removing, throwing, or destroying the can. Before each trial, the lid was placed in one of three positions: on top of the trashcan, to the side of the trashcan, or in front of the trashcan.

In the top condition, the lid rested on top of the trashcan (Figure 2). At SFDK, with the curved lid, it was placed upside down with the curved side going into the trashcan so that it was very easy for the elephants to remove. At MZ, with the flat lid, the lid was still upside down (so that it did not lock into place), but was more difficult for the elephants to remove when lifting, although it could easily be slid off to the side.

In the side condition, the lid was placed 0.3048 meters to the right of the trashcan. The lid was placed upside down (with the curved side of the lid on the ground at SFDK).

In the front condition, the lid was placed upside down 0.3048 meters in front of the trashcan (Figure 1). At SFDK, the curved side of the lid was on the ground, so that the orientation of the lid was the same whether it was on or off the trashcan. The flat lid was also placed upside down on the ground at MZ.

Each elephant participated in 15 consecutive trials. Five trials for each of the three lid conditions were embedded across the 15 trials in a predetermined random order. The elephants were randomly assigned to one of two treatment groups: A or B. Group A started with the lid on top of the trashcan, while group B started with the lid in front of the trashcan.
Experimental Procedure

*Experimental trials.*

During trials at SFDK, the elephant participating in the experiment was walked from the exhibit, through a fenced gate, to a back area where the experiment was performed and unable to be seen by the other elephants. The other elephants remained on exhibit, omitting the opportunity for observational learning. For each trial, the elephant to be tested was walked to the exhibit side through the fence she had come through to control for a number of potential biases, including the type of reinforcement provided, experimenter interaction with the lid, and a preview of the next condition. Once these pieces were completed, the elephant was returned to the area where the trashcan was presented.

Trials at MZ were conducted in a similar fashion. Since the facility setup was slightly different, the elephant performing the trials remained in the corral until the experiment was complete, while the other elephants were in a much larger open corral with other hoofstock. In between trials, the keeper walked the elephant to the opposite side of the empty corral, keeping the elephant’s focus away from where the experiment was being reset for the next trial. Due to social paradigms, Christie and Paula were in the corral at the same time to avoid anxiety due to being separated. For their trials, two trainers were used. While one elephant was engaged in a trial, the keeper kept the other focused away from the trashcan (in a similar fashion to when the trials were being reset), so that observational learning could not take place. Additionally, in order to further avoid confounds, they were not in the same treatment groups, so the lid placement was not the
same for their trials. The other three elephants were in the corral one at a time during their experimental sessions.

One piece of produce (usually sweet potato), from the approved daily diet, was placed in the bottom of the trashcan as reinforcement for each trial. Elephants did not receive any training or direction from their trainer on how to interact with the trashcan. Since elephants have an excellent sense of smell, the keepers were confident that the elephants would be able to detect the food, even though it was in an opaque container. Consequently, the elephants’ natural foraging behavior, as well as their intrinsic motivation to acquire the food, and their inquisitive nature lead the elephants to naturally interact with and manipulate the trashcan to obtain the food without training. If the elephant did not direct his or her attention to the trashcan after 1 minute, additional food reinforcement was added to the trashcan, following the same procedure as resetting a trial.

The area where the testing occurred was in a location that the elephants did not usually access. Thus, if the elephants began to investigate the surroundings instead of the trashcan, the trainer would redirect the elephant’s attention back to the trashcan. The trainer did not, however, give any direction to the elephant as to how the trashcan or the lid should be manipulated.

If the elephant tried to take the lid with him or her when leaving the experiment site, the trainer would signal the elephant to drop the lid, or if the elephant tried to hand the trainer the lid, the trainer would knock the lid to the ground and not accept the lid or reinforce this undesired behavior. Since the elephants often engaged in training sessions when their trainers were around, reinforcing behaviors would lead to an increase in the
frequency with which they occurred. Manipulating the lid in conjunction with the trainer was not the desired behavior, thus the trainer did not reinforce this behavior or the elephant would have perceived it to be a desired behavior and would have done it more frequently.

To minimize the possible influence of the trainer’s gaze, after leading the elephant to the trashcan area, the trainer would stay back so that he or she was positioned to the side of the elephant upon the start of the trial and remained about halfway back on the elephant so that he or she was not directly in his or her line of vision when facing the trashcan (see Figure 1 and Figure 2).

The trial was completed once the elephant retrieved the food reward from inside the trashcan and placed it in his or her mouth. Consequently, it was not necessary to touch or manipulate the lid to complete a trial. Following the end of the trial, the elephant was walked back to the exhibit (SFDK) or to the opposite side of the corral (MZ) so that the next trial could be set. This process was used until all 15 trials were completed. A session usually lasted 30 minutes or less.

Coding of trials.

Each trial was coded to determine three main points of interest. First, it was determined if the elephant touched the lid before acquiring the food reward from inside the trashcan. Second, the latency from the start of the trial to when the elephant placed the food in his or her mouth was calculated. Third, whether or not the elephant completed the trial (by acquiring the food) was determined.

Lid touches for the side and front condition were coded using a touch no-touch count. If an elephant did not touch the lid, it resulted in a value of zero.
Latency, or time that it took to acquire the reward, was also be measured. This time started when the elephant reached a set point, about 138 inches away from the trashcan. This distance was far enough away that the elephant could not reach the trashcan or lid with his or her trunk until stepping closer. The trial ended when the elephant placed the food in his or her mouth. If the elephant was learning as the trials went on, it would be expected that the numbers would decrease as the trials continue. It was also anticipated that the trials with the lid on top would have a higher latency than the side or front conditions, due to the removal of the lid.

Statistical Analyses.

A Pearson’s chi-square test of independence was used to determine if there was a relationship between two of the lid treatments (top first or front first) and the number of trials in which the elephant touched the lid before acquiring food. It was predicted that elephants who first encountered the lid-on-top condition would then associate moving the lid with acquiring food, consequently touching the lid more in the other conditions, where it was not a necessary step to acquiring the food. A second Pearson’s chi-square test of independence was performed for lid position and its relationship to whether the lid was touched before obtaining food. Since the elephant had to touch the lid in the top trials, it was hypothesized that the top condition would have more lid touches overall than the other two lid positions. Using only the first trial for each elephant, a final chi-square goodness of fit test determined if there was a significant relationship between the number of lid touches and lid placement (front or top). This analysis allowed for a comparison of the elephant’s behavior in the initial exposure to the front or top paradigm. Since the elephants had to touch the lid to acquire the food in the top condition, this allowed for a
comparison to see if there was initially a significant relationship in behavior between the two conditions.

To determine if the position of the lid, either front or top, significantly influenced latency to obtain the reward on the first trial, an independent-samples t-test was conducted. It was hypothesized that the obstacle of having the lid on top of the trashcan would lead to significantly longer latencies on the first trial.

A one-way repeated measures ANOVA was used to determine if there was an effect of lid position on the latency to retrieve food. An average latency was used for each of the three lid positions per elephant to minimize the within subject variation and increase power. It was predicted that the latency to retrieve the food would be greatest for top trials.

*Intercoder Reliability.*

Intercoder reliability for whether or not the elephant touched the lid on the initial trial was performed for 20% of the trials, which were chosen at random. Cohen’s (1960) kappa formula, \( \kappa = (p_o - p_e)/(1 - p_e) \), wherein \( p_o \) = observer agreement and \( p_e \) = agreement expected by chance, was used to calculate reliability, which was set for 80% agreement, between two independent coders. The correlation for the latency of the trials, coded by two independent coders, was also determined. Lid touches before acquiring the food reward were in complete agreement between coders, \( \kappa = 1, \ p < .00 \). Latencies were highly correlated, \( r (25) = 1, \ p < .00 \). As a result, all reliability criteria were met.
**Figure 1.** Front trial at SFDK.

This figure shows the experimental setup at SFDK in the back area behind the exhibit. This trial had the lid in the front placement.

**Figure 2.** Top trial at MZ.

This figure shows the experimental setup at MZ in the elephant yard. This trial had the lid in the top placement.
CHAPTER III - RESULTS

Each elephant successfully acquired the food from the trashcan in all completed trials, 15 per elephant.

Two elephants, Malika (top placement first) and Valerie (front placement first), only touched the lid in the lid-on-top trials. All of the other elephants touched the lid on trials when the lid was in the side or front position, and a lid touch was not necessary to acquire the food. Elephants in the top first treatment touched the lid before acquiring the food on 53.3% of the 75 total trials (five elephants, 15 trials per elephant). Elephants in the front first treatment touched the lid before acquiring food on 46.6% of the 58 total trials (four elephants, 15 trials per elephant, two trials excluded due to a recording error). The chi-square goodness of fit test indicated that the elephants were equally likely to touch or not touch the lid before acquiring food, regardless of the position treatment each elephant experienced first (top or side), $x^2(1, N = 133) = 0.60, p = 0.44$.

The 3x2 chi-square test of independence revealed that the percentage of trials in which elephants touched the lid before acquiring food versus those in which the elephant did not touch the lid varied based on the lid placement, $x^2(2, N = 133) = 69.96, p < 0.001$. Elephants touched the lid 100% of the time in the top position, 13.6% of the time in the side position, and 37.8% of the time in the front position. The adjusted standardized residuals were calculated to see where the number of touches differed significantly from the expected number of touches due to chance. This value was significant for all three lid positions (top adjusted residual = 8, side adjusted residual = 6, front adjusted residual = 2.1).
The chi-square test of independence conducted to examine the first trial of the top and front lid treatments found that the elephants did not have an equal percentage of lid touch and non-lid touch trials. Elephants touched the lid 100% of the time on the first trial in the lid on top treatment and 50% of the time on the first trial in the lid-in-front treatment. However, likely due to the small number of trials, the results of the test did not meet an alpha of .05, $\chi^2(1, N = 9) = 3.214, p = 0.07$.

An independent-samples t-test was conducted to determine if the lid placement, either front or top, impacted the latency of the first trial. There was not a significant difference between the top ($M = 93.4$ s, $SD = 60.71$ s) and the front ($M = 69.75$ s, $SD = 98.73$ s) conditions, $t(7) = 0.45, p = 0.67$. Latency to retrieve food was not influenced by the starting position of the lid when experienced as the first trial.

A one-way repeated measures ANOVA was conducted to compare the effects of lid position on mean latency from the start of the trial to when the elephant placed the food reward in his or her mouth in the front, side, and top conditions. There was a statistically significant effect for lid position, Wilk’s Lambda = 0.023, $F(2, 16) = 8.25, p = .003, \eta_p^2 = 0.51$. Post hoc tests using a Bonferroni correction revealed that the mean latency for the top condition ($M = 38.29$, $SD = 24.89$) was significantly longer than the side condition ($M = 9.38$, $SD = 3.34$), $p = 0.019$. On the other hand, the mean latency for the top and front ($M = 14.13$, $SD = 14.73$) and the front and side lid placements were not significantly different from each other.
CHAPTER IV – DISCUSSION

Causal reasoning is difficult to discern in nonverbal humans, and even more challenging to unravel in nonverbal non-human animals (Blaisdell et al., 2006; Gopnik & Glymour, 2002; Gopnik & Sobel, 2000; Gopnik et al., 2001; Gopnik et al., 2004; Kuczaj et al., 2009; Mulcahy & Call, 2006; Nissani, 2006; Premack & Premack, 1994; Povinelli, 2000; Seed et al., 2006; Silva et al., 2005; Sobel & Kirkham, 2006, Sobel & Krikham, 2007; Taylor et al., 2009; Tebbich & Bshary, 2004; Visalberghi & Limongelli, 1994; Visalberghi & Tomasello, 1998; Zuberbuhler, 2000). One of the biggest challenges in this task is finding a paradigm that not only reliably investigates the desired question, but also does so with an evaluation and a metric that is meaningful and biologically relevant to the animal being tested. If an animal is not interested in the paradigm, this can easily lend itself to the conclusion that the animal does not understand the paradigm, when instead the animal is simply not engaged in the task. Likewise, factors such as previous experience and training can shape an animal’s future behavior and interaction with an experimental setup (Pavlov & Anrep, 1927; Skinner, 1983; Thorndike, 1927). As a result, replication studies that support previous research are one way to contribute to establishing the validity of a concept.

The purpose of the replication study was to determine if the elephants would produce similar behavioral interactions with a trashcan lid obstacle-removal paradigm to those found by Nissani (2006). The original Nissani (2006) study concluded that elephants could not causally comprehend that the lid was only an obstacle to the food in a bucket when the lid was placed on top of the container that held the food, and not when it was placed to the side of it. Elephants in Nissani’s (2006) study consistently tossed the
lid before attempting to retrieve their food reward from the bucket. Even when the lid was off to the side of the bucket, and did not serve as an obstacle to the food, the elephants would throw the lid more often than not. Consequently, the present replication aimed to determine if the placement of the lid reliably affected the elephants’ behavior. Using two different lid treatments: one that started with the lid on top of the trashcan, and one that started with the lid in front of the trashcan (relative to the elephant’s position), the effects of the lid placement were examined in regards to the elephants’ elicited response to the paradigm on the first trial, the elephants’ lid touching behavior, and the length of the interaction with the lid and trashcan before acquiring the food reward.

Two of the nine elephants, who were given different treatments and located at separate facilities and thus unable to influence one another, only touched the lid on the top trials, ignoring it in the other two conditions. This qualitative difference in performance counters Nissani’s (2006) results, which argued against the elephants’ propensity toward causal knowledge, as none of his elephants ignored the lid on all non-top lid placement trials. Since elephants housed at two separate facilities independently produced these results, which were never seen in the Nissani (2006) study, it is likely not due to chance. The behavioral patterns of the other seven elephants provide additional evidence for elephants having some level of understanding of the causal reasoning needed to solve this task.

Both elephants at SFDK who received the lid-on-top first treatment displayed a behavior where they placed the lid back on the trashcan after eating the food. This behavior was only present at SFDK and only for the elephants who received the top first treatment, suggesting that the lid-first placement influenced the elephants’ interactions
with the lid in subsequent trials. This behavior could be interpreted as superstitious, but Kellogg (1949) and Skinner (1948) argued that such an interpretation would be too anthropomorphic. Instead, Kellogg (1949) and Skinner (1948) argued that animal behavior is more often the result of the consequences that follow that behavior, which aligns with a behaviorist view of learning (Skinner, 1983; Thorndike, 1927). The behaviorist model of learning dictates that if there were positive consequences for the behavior, the animal would be more likely to repeat the behavior and repeat it more often. Accordingly, it is possible that the positive reward (food) retrieved after removing the lid from the trashcan lead the elephants to reproduce lid manipulation behaviors on subsequent trials, perhaps expecting similar consequences. Moreover, the temporal pairing of the lid being removed right before the food was acquired could have created a spatio-temporal relationship, suggestive of associative learning, instead of causal reasoning (Visalberghi, & Tomasello, 1998). Nissani (2006) also suggested that the elephants in his study might have been engaging in associative learning instead of causal reasoning.

Since the elephants in the Nissani (2006) study received extensive training, this had the potential to create not only a temporal association between throwing the lid and acquiring a food reward, but also, according to behaviorist models of learning, increasing the probability that the behavior would be reproduced due to the reinforcement potentially being associated with the lid movement (Skinner, 1983; Thorndike, 1927). Furthermore, this learned association would be difficult to extinguish in just a few trials. In fact, in order for extinction of the behavior to occur, removing the lid would have to no longer be reinforced (Nevin, 2012). Since the lid tossing behavior in the Nissani (2006)
study continued to be reinforced on all top trials (as well as in other concurrent experiments that were happening between sessions), this variable reinforcement schedule made it highly unlikely that the elephants would stop engaging in the lid tossing behavior regardless of their understanding or lack of understanding of the obstacle-removal lid paradigm (Nevin, 2012). Likewise, it is possible that the lid manipulations after receiving the food reward, in the current study, were a similarly learned association due to the temporal relationship between removing the lid and acquiring the food.

While the elephants in the current study did not receive any training or pre-training on how to interact with the paradigm, one of the things causal reasoning is contingent upon is having experience (Premack & Premack, 1994; Visalberghi & Tomasello, 1998). Only through experience or previous knowledge can a cause and effect relationship be established, but this causal connection can also take time to discover. This control created a fundamental difference between the Nissani (2006) study and the present study. The elephants in the Nissani study already had an established relationship with the paradigm before they started their experimental trials, while the SFDK and MZ elephants were establishing that relationship in their initial trials. This means that the training trials for the elephants in the Nissani (2006) study, which in theory were the elephants’ only exposure to this paradigm, generated a relationship between throwing the lid and receiving a reward. Consequently, the connection between throwing the lid and receiving a reward would have to be extinguished to correctly complete the trials in the other lid positions (Nevin, 2012). Furthermore, in the Nissani (2006) study there was no clear signal to the elephants that the desired interaction with the paradigm should change when the trials shifted from
training to experimental. Hence it was only from their previous experiences, the training trials, that the Nissani (2006) elephants knew how to interact with the paradigm. Accordingly, it is not surprising that they continued to engage in the lid throwing behavior which was previously trained and continued to be reinforced. The current study, however, eliminated the potential confound of training and did not train the elephants on interacting with the apparatus at all (see Table 2 for a summary of differences).

Behavioral differences, such as placing the lid back on the trashcan at SFDK, suggested that there were qualitative behavioral trends that disagreed with the Nissani study (2006). First, the overall lid placement treatments resulted in different behaviors displayed by the elephants (e.g., placing the lid back on the trashcan or only touching the lid on top trials). Second, even though there might not have been enough trials for the lid position treatment to have a statistically significant impact on behavior on the first trial, the lid position significantly influenced whether or not the elephants touched the lid before retrieving the food reward overall. Elephants touched the lid 100% of the time in the top position, 13.6% of the time in the side position, and 37.8 % of the time in the front position. One interpretation of this significant difference in interactions based on lid position is that the elephants were more likely to touch the lid on the way to the trashcan, in the front position, than they were when the lid was off to the side. This differential responding suggests that unlike in the Nissani study (2006), the current elephants in this study had different behavior patterns depending on the position of the lid. This behavioral difference could be indicative of a level of causal reasoning in this paradigm as the elephants usually did not go out of their way to touch the lid. It is possible that since, in the front position, the lid was between the elephant and the food,
the elephant was more likely to touch the lid than when the lid was off to the side. This would support the argument that it was not a lack of causal understanding that led the elephants to interact with the lid, as they did not frequently go to the side of the trashcan to touch the lid before acquiring the food. If the elephants did not understand that the food was not always under the lid, as Nissani (2006) suggested, then it would be expected that the elephants would have interacted with the lid in the side position as much as they did when it was in the front position. Since this was not the case, it supports the theory that the elephants could have had some understanding of the paradigm. An additional study would need to be performed to conclusively determine if the elephants changed their behavior due to an understanding of the paradigm or if their behavioral differences were due to the lid being slightly more out of sight and out of mind in the side placement condition.

Elephants touched the lid 100% of the time on their first exposure to the paradigm when the lid was on the top of the trashcan, but only 50% of the time when the lid was in the front. While this result did not meet the established alpha of .05 (p = .07), a qualitative interpretation of the data suggested that the elephants seemed naturally curious about the lid since they interacted with it 50% of the time, without understanding the paradigm. This could also be explained by the necessity for the elephants to gain experience interacting with the paradigm to establish a causal relationship (Premack & Premack, 1994).

Discerning level of interest is difficult in a non-verbal animal. The most frequently used measure, as was seen with many other comparative studies (Gopnik et al., 2001; Leslie, 1982; Leslie & Keeble, 1987; Oakes & Cohen, 1990; Schlottmann &
Surian, 1999; Sobel & Kirkham, 2004; Sobel & Kirkham, 2006) is attention, or time spent interacting or gazing at a stimulus. In this instance, the easiest measure of attention was the time that it took the elephant to acquire the food reward, or latency. If the elephant was focused only on the task at hand, it should take less time to acquire the food than if the elephant is paying attention to the lid before acquiring the food. Consequently, the measure of latency can be considered a reflection of the elephant’s attention to both the lid and acquiring the food reward.

It was predicted that the top trials would have a longer latency than the other two lid positions as the top position required the elephant to pay attention to and remove the lid before acquiring the food. In contrast, the side and front conditions did not necessitate an interaction or attention to the lid to obtain the food reward, and should, consequently, have shorter latencies. While there was no significant difference in the latency of the first trial exclusively, based on lid position (front or top), a significant difference in latency was found between the top and the side lid placements overall. This significant finding corresponds with the percentage of trials in which the elephants interacted with the lid: 100% of the time in the top position, 13.6% of the time in the side position, and 37.8% of the time in the front position. If latency is a reflection of attention, these results supported the prediction that when the lid was off to the side it did not draw the elephant’s attention, or get touched, as much as when the lid was on top of or in front of the trashcan. Again, additional research would be needed to know if the attention given to the lid in the front position was due to the elephants spatially encountering the lid on the way to the trashcan, or if there is another explanation for this behavior.
One of the limitations to this interpretation is that in some of the trials, the elephants appeared to be less motivated to acquire the food. They took significantly more time to explore their surroundings before drawing their attention back to the trashcan. On a couple occasions, the elephant was walked away from the trashcan and additional food was added to the trashcan in an attempt to encourage the elephant to participate. Having a larger number of trials to draw from would help normalize these outliers, although a larger number of trials per session could also put the elephants at risk of losing interest in the paradigm altogether.

Overall, the results of this study did not support those of Nissani (2006). Behavioral differences with regards to lid touches and trial latencies were not the same in all three lid positions. These results are far more suggestive of the potential for a causal understanding of the paradigm than were those from Nissani (2006). Additionally, the current study suggests that the results of the Nissani (2006) study were likely more of a reflection of temporal associations and training than they are a reflection of causal reasoning (Table 2). By omitting any training of the paradigm or exposure to the paradigm before the trials, the elephants in the current study were able to interact with the paradigm in an unbiased way and did not show the same patterns of behavior as those in the Nissani (2006) study.
Another question that arises is whether or not the Nissani (2006) paradigm really investigated causal reasoning or tool use. Aside from potentially reflecting training, and consequently associative learning more than causal reasoning, there is also the question of the cause and effect relationship Premack and Premack (1994) identified as being essential to any causal reasoning paradigm. Nissani (2006) argued that the elephants did not use tools, but the Nissani Study refers to the paradigm as a tool use task. According to the Schumaker et al. (2011) definition, no external object was used or manipulated to achieve the desired outcome. Initial training could have created the relationship between lid tossing and the food reward. They could, but behavioral differences in the 3 lid conditions suggest a potential causal knowledge about removing the lid obstacle to acquire food.
not have causal reasoning because he believed that they threw the lid, even when it was 
not necessary, to look for food under the lid. This argument does not account for an 
elephant’s strong sense of olfaction. Elephants, with their specialized scent glands, have 
been shown to have an extremely keen sense of smell (Rizanovic, Amundin, & Laska, 
2013). Asian elephants out-performed humans, among other species, in an olfactory 
discrimination test between structurally related odorants (Rizanovic et al., 2013). With 
such a profound sense of smell, it seems improbable that the elephants were not 
immediately able to discern both that a food reward was present and exactly where that 
food reward was. In fact, it is likely due to their sense of smell and detection of the food 
reward that the elephants in the current study were able to successfully interact with the 
paradigm without any training. The elephants explored the apparatus, but ultimately did 
not appear to have any difficulty discerning that the goal of the task was to obtain the 
food reward in the trashcan, which they could clearly detect, as the elephants were blind 
to the placement of the reward. Consequently, if the elephants knew where the food was, 
there has to be another reason for them to interact with the lid of the trashcan.

Performing trials where an elephant interacted with the apparatus, without any 
food reinforcement, training, or previous experience with a trashcan or a lid, might help 
reveal the elephants’ tendency toward the lid touching behavior. A different paradigm, 
which more clearly looks at a causal relationship would also paint a clearer picture of the 
elephants’ causal understanding.

In Nissani’s study (2006) he alluded that he wanted his research to investigate the 
similarities between the elephants and the lid paradigm and Thorndike’s cats (Thorndike, 
1927). Assuming that there is a causal relationship in the Nissani study (2006), it would
be that touching or throwing the lid when it was not in the top position caused the elephant to no longer get the food reward, but throwing the lid when it was still in the top position still resulted in a reward. Consequently, unlike Thorndike’s paradigm (1927), the cause and effect relationship was situationally dependent. If the lid was in one position it required one behavior, and if it was in another it required a different behavior to acquire the food reward. Furthermore, in order to look at the extinction curve of the behavior, the stimulus response relationship that is to be extinguished can no longer be rewarded (Nevin, 2012; Thorndike, 1927). The paradigms are too dissimilar methodologically to investigate the same hypotheses. Successful completion of the Nissani paradigm was not contingent upon any knowledge of the cause and effect relationship needed to differentiate between causal reasoning and associative learning (Premack & Premack, 1994). If the elephants in the Nissani (2006) study had retrieved the food without touching the lid on the side trials, then they never would have been exposed to the causal relationship that they were supposed to be learning. Consequently, the Nissani (2006) paradigm was not well-suited for investigating causal reasoning.

Furthermore, using the operational definition of tool use set by Shumaker et al. (2011), it is unclear why Nissani (2006) refers to his paradigm as being a tool-use task:

The external employment of an unattached or manipulable attached environmental object to alter more effectively the form, position, or condition of another object, another organism, or the user itself, when the user holds and directly manipulates the tool during or prior to use and is responsible for the proper and effective orientation of the tool. (p. 5)
The elephant’s trunk is not considered a tool, as it is an attached appendage. Likewise, the lid is not a tool since it was not used to alter anything else. The lid also was not required to complete 2/3 of the trials. Since there was no additional tool utilized in the paradigm, Nissani’s (2006) paradigm does not meet the criteria for a tool-use study. Consequently, to draw conclusions about tool-use in elephants, a paradigm that incorporates an external object to manipulate something would be required.

In conclusion, the conundrum to investigating the capacity for causal reasoning in elephants lies primarily in finding a paradigm that reliably investigates the question using biologically relevant methodology for the elephants. For example, instead of using a trashcan and a lid, if a hollowed tree stump was covered by a removable branch with foliage, this would make the apparatus more similar to something that the elephant would naturally encounter when foraging. Making the paradigm more naturalistic might also decrease the novelty of the items and allow the elephants to be more focused on retrieving the food and less focused on interacting with novel items. To date, the Nissani study (2006) is the only published research that claims to address causal reasoning in elephants, and the validity of those results were questioned by the current study. Rather than negating the capacity for causal reasoning in elephants as Nissani concluded, the current study suggests otherwise. At a time when laws about elephant management are changing, it is imperative that researchers paint the clearest picture possible about the cognitive abilities of elephants so that those animals that are in human care can have the best environment and most enriched lives possible. Likewise, from an evolutionary standpoint, knowing more about the cognitive abilities of elephants can only help to understand the origins of human cognition and the evolutionary advantage that these
abilities have provided. Furthermore, since there is no singular validated test for causal reasoning, and the most commonly used test, the trap-tube test, consistently yields mixed results, there is the need for a new paradigm that could produce more conclusive results (Blaisdell et al., 2006; Gopnik & Glymour, 2002; Gopnik & Sobel, 2000; Gopnik et al., 2001; Gopnik et al., 2004; Kuczaj et al., 2009; Mulcahy & Call, 2006; Nissani, 2006; Premack & Premack, 1994; Povinelli, 2000; Seed et al., 2006; Silva et al., 2005; Sobel & Kirkham, 2006, Sobel & Krikham, 2007; Taylor et al., 2009; Tebbich & Bshary, 2004; Visalberghi & Limongelli, 1994; Visalberghi & Tomasello, 1998; Zuberbuhler, 2000). Elephants with their large brains and high encephalization quotients (Roth & Dicke, 2005) are logical candidates for this research as soon as the proper paradigm is in place.
APPENDIX A – IRB Approval Letter

TO: Dr. Karen Coats, Dean of the Graduate School
FROM: Dr. Sam Bruton, Director, Office of Research Integrity
COPY: Alen Hajnal
Dr. Jake Schaefer
Dr. Joe Olen, Chair Psychology
DATE: 11 May 2017
SUBJECT: Beri Brown Research

Beri Brown, a graduate student in Psychology, is conducting thesis research under the direction of Dr. Alem Hajnal, and this research entails review/analysis of video-taped behavior of captive elephants at Six Flags Discovery Kingdom and Monterey Zoo.

Although the Graduate School requires Institutional Animal Care and Use Committee (IACUC) approval of all research projects involving vertebrate animals, the federal regulatory agencies that oversee the University’s Animal Care Program and IACUC activities do not require an approved IACUC protocol for observational activities that do not alter or influence the activity of the animals, impact the health of safety of personnel, or impact the animal’s environment. The project in question involves analysis of archival video and therefore does not alter or influence the activity of the animals. Neither does it impact the health or safety of personnel conducting the study or the animals’ environment.

Please accept this letter in lieu of IACUC approval of Brown’s research protocol.

Thank you for your consideration of this request.

Sam Bruton
Director, ORI
APPENDIX B – SFDK Approval Letter

Vanessa Fravel DVM  
Six Flags Discovery Kingdom  
1001 Fairgrounds Drive  
Vallejo, CA 94589

April 23, 2017

Beri Brown  
The University of Southern Mississippi  
225 South Heights Blvd #1413  
Houston, TX 77007

Dear Ms. Brown,

I am writing to inform you that your research proposal has been accepted by the Animal Care and Use Committee at Six Flags Discovery Kingdom. You may use the data you acquired from our elephants for your Master’s thesis. We ask that you send us a copy of your work prior to publication in a peer reviewed journal in order to be reviewed by the committee.

If you have any questions or concerns please contact me.

We look forward to reading your work.

Sincerely,

Vanessa Fravel DVM
APPENDIX C - MZ Approval Letter

4/25/2017

To whom it may concern:

Beri Brown is granted permission by the board of directors of Elephants of Africa Rescue Society (EARS) to use the elephant data she obtained in 2011 at our facility while accomplishing her replication study of Nissani’s 2006 lid paradigm study. It is our understanding that the information she obtained will be used to complete her Master’s Thesis for the University of Southern Mississippi.

We are proud to have been able to assist Beri as well as to be a part of this study. Beri has our permission to use and/or share the information she obtained in her study in any form or fashion that she feels will responsibly better the lives of all elephants, wild and captive.

We look forward to any and all positive influences that might result from her efforts.

Sincerely,

Charlie Sammut
CFO, EARS

E.A.R.S. Mission Statement: Elephants of Africa Rescue Society is dedicated to supporting projects aimed at securing a safe habitat for wild African animals and dedicated to providing a sanctuary for their captive cousins with a special focus on elephants. We believe that by educating the public about the problems facing both captive and wild animals we can increase public awareness and support.
REFERENCES


mechanisms in very young children: Two, three, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental psychology*, 37, 620-629.


Rizanovic, A., Amundin, M., & Laska, M. (2013). Olfactory discrimination ability of
Asian elephants (elephas maximas) for structurally related odorants. *Chemical Senses*, 38(2), 107-118.


