Surprise-induced exploration as a tool for learning: A comparative approach with human infants and non-human primates

Citation

Walco, Emily Rachel. 2022. Surprise-induced exploration as a tool for learning: A comparative approach with human infants and non-human primates. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Human Evolutionary Biology

have examined a dissertation entitled

“Surprise-induced exploration as a tool for learning: A comparative approach with human infants and non-human primates”

presented by Emily Rachel Walco

candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Signature Martin Surbeck
Martin Surbeck (Aug 23, 2022 21:06 GMT+2)

Typed name: Prof. Martin Surbeck, Harvard HEB

Signature Erin E Hecht

Typed name: Prof. Erin Hecht, Harvard HEB

Signature

Typed name: Prof. Elizabeth Spelke, Harvard Dept. of Psychology

Signature Christopher Krupenye

Typed name: Prof. Christopher Krupenye, Johns Hopkins University

Date: 18 August 2022

Surprise-induced exploration as a tool for learning: A comparative approach with human infants and non-human primates

A dissertation by

Emily Rachel Walco

Human Evolutionary Biology

Harvard University

Cambridge, MA

August 2022

Dissertation advisor: Prof. Elizabeth Spelke Emily Rachel Walco

Surprise-induced exploration as a tool for learning: A comparative approach with human infants and non-human primates

Abstract

Upon witnessing a surprising event, young humans will often explore the target of that surprise, appearing to seek an explanation (eg: Bonawitz et al., 2012; Chandler & Lalonde, 1994; Perez & Feigenson, 2022; Stahl & Feigenson, 2015; van Schijndel et al., 2015). Surprise-induced exploration could serve to reveal otherwise opaque causal mechanisms, helping to scaffold causal reasoning so that learners can begin to think abstractly about how entities in the world relate. Humans seem uniquely capable of this type of reasoning. Could it be that humans' tendency to seek explanations sets them apart from even their closest living relatives? In this thesis, I aim to shed light on this question by conducting comparative research with humans and non-human primates. In Chapter 2, I assess looking time as an indicator of surprise, one of the primary tools that has been used to assess non-verbal individuals' expectations. I argue that this method, while still extremely valuable and worthwhile, may not be as flexible and robust a measure as it is often taken to be, because neither infants nor monkeys looked longer at events that past research tells us they should find surprising (Baillargeon, 1987, 1995; Spelke et al., 1992). In Chapter 3, I present a paradigm for assessing individuals' expectations about how objects act and interact in the world, demonstrating that bonobos, one of humans' closest living relatives, can use principles of object dynamics to locate hidden items. And in Chapter 4, I provide evidence that bonobos do not preferentially explore objects that are unexpected, suggesting that they do not seek explanations for surprising events. The research presented in this thesis ultimately provides evidence that humans and primates do not seem to differ in their underlying expectations about how objects interact in the world, but humans may be unique in their propensity to harness instances where these expectations are violated to scaffold more abstract causal reasoning.

Table of Contents

List of Figures

List of Tables

Acknowledgements

The work presented in this thesis was made possible by the generous support of the Mind Brain Behavior Initiative, Foundations of Human Behavior, the Center for Brains, Minds, & Machines, the Cora Du Bois Fellowship, and the Human Evolutionary Biology Department.

I am grateful to so many people who have shaped my path to and through this PhD journey. I certainly could not have done this alone. To all those, both human and non-human, who made this work possible, thank you.

First, to my committee, thank you all for your feedback, advice, and encouragement. To my primary advisor, Elizabeth Spelke, thank you so much for welcoming me into your lab and for helping me find my footing when I needed it. I have learned a tremendous amount from you, and you have made me feel like I'm not a *total* stranger in developmental psychology. It has been so fun playing with these ideas with you, and I sincerely appreciate your encouragement and continued excitement and curiosity, and allowing me the flexibility to try things out, with babies and apes alike. Martin Surbeck, thank you for providing me a home in HEB. It truly meant so much to me when I returned from my semester in Des Moines in 2019 and was immediately welcomed into your lab group. The advice, community, and company of fellow primatologists that you and your lab provided were invaluable. Erin Hecht, thank you for your continuous support and encouragement. There have been so many times that I have felt reassured after talking to you. You have also been an advocate for me, and for that I am so grateful. Last but certainly not least, Chris Krupenye, thank you for helping me navigate the ins and outs of conducting research with apes. You made me feel like my work was exciting, and valuable, and doable, and you helped me through all the tricky steps of making it happen. Your words of encouragement and your enthusiasm about the research kept me going.

I also owe a great deal of thanks to members of the Spelke lab. Shari Liu and Ashley Thomas, your mentorship throughout my PhD has helped shape me into the scientist that I am. I have come to each of you more times than I can count with questions, concerns, and ideas to think through, and you have made me feel heard and supported each and every time. I look up to both of you as researchers and mentors, and I cannot thank you enough for being such wonderful role models. To my fellow graduate students, Brandon Woo and Akshita Srinivasan, thank you both for your encouragement, commiseration, and comradery. And to the Spelke lab managers, Bill Pepe and Cristina Sarmiento, your efforts have made this research possible. Thank you for all that you do.

Thank you as well to the Human Evolutionary Biology Department, for providing me a space to learn and grow and challenge myself. Erin Wessling, Liran Samuni, and Rachna Reddy - each of you is so impressive and insightful, and I cannot thank you enough not only for your endless willingness to help and for your mentorship, but also for your friendship. It shaped my experience in this department to have you all here. I am also beyond grateful for the friendships I have made through this department. Evvie Jagoda, Katia Chadaideh, and Laura Lewis, we became friends within my first month in the department, and it profoundly shaped my experience here from the start. I genuinely do not think I could have persevered through everything that this PhD has thrown my way without your friendship and guidance. I am also so appreciative for the Diversity, Inclusion, and Belonging Committee. This group created such a warm and supportive space that made me feel like we were striving for better and gave me hope for the future. Richard Wrangham, thank you so much for helping me to ground this work in the context of human evolution. Your mentorship early in my PhD was pivotal to the development of this research and to my development as a scholar. To the HEB administrative staff, especially Meg Lynch, Mallory McCoy, Monica Oyama, and Natalie McKenna, thank you so much for your tireless efforts to keep the department going and for helping me to navigate all the ins and outs of this PhD.

There are so many individuals who supported my research efforts. First to the Ape Initiative staff, Jared Taglialatela, Amanda Epping, Chelsea Cox, Emily Cain, Anna Olson, Calle Uerling, and Haley Holmes - thank you not only for all of the time and effort you dedicated to making my studies happen, but thank you as well for providing such thoughtful and loving care to those wonderful bonobos. To each of the research assistants who have dedicated time to my research and have helped me grow as a mentor, thank you. In particular, Hayoung Chang, Cameron Calderwood, and Vanessa Kudrnova, thank you for the countless hours you spent tediously coding where babies and primates look and point. Kris Brewer, thank you for helping me create stimuli more times than I'd like to count – you made the magic happen. And to Steve Worthington, I appreciate all of your advice and guidance on conducting the analyses presented in this thesis. Thank you also to the many infants who I have tried to surprise, and of course to their parents, as well.

I have so much gratitude for Jason Mitchell for showing me a path when I couldn't find one and encouraging me in my career as a scientist. I can truly say that I am not sure I'd have made it here without your guidance and support.

I owe so much of my career as a scientist and a primatologist to Anna Weyher. You took a chance on me about one month into my freshman year of college when I expressed interest in doing research with primates. I learned a tremendous amount from you in so many ways, and I cannot thank you enough for the opportunities you provided me. I came to WashU with a dream and a passion, and you realized that dream and helped that passion grow. Thank you for sharing your wisdom and mentorship, your baboons, and your friendship. You shaped my career and who I am.

I also need to recognize and thank all of those who supported me and believed in me outside of my professional setting. First, to my parents, your unconditional love and encouragement have meant the world to me. There really aren't enough words to express my gratitude. You have made me feel seen, valued, and supported for as long as I can remember. From reading my Animal Encyclopedia with me innumerable times, to driving me to the Bronx Zoo from New Jersey every day for weeks, to visiting me at my field sites, and so much more, you have fostered my passion for animals into a career. You make me feel capable and valuable, and I owe so much of who I am to the two of you. To my brothers, Jeremy and Dan, you have always given me very big shoes to fill, and I will spend my life trying to live up to the impressive and intelligent people that you both are. I feel so lucky to have you as my big brothers. To my sister-in-law Melissa, cliché as it might sound, you are the sister I never had. Thank you for sharing your love of animals and nature with me; your wonder for the natural world amplifies my own. To my nieces and nephew, Evey, Teddy, and Penny, I am so fortunate to be your aunt. Your curiosity and zest have fueled my interest in how we come to make sense of the world, and you've even helped me work out some of my methods. Each one of you is warm, and funny, and spunky, and loving. To each member of my family, I love you so much. Thank you.

To all of my friends, those from way back, and those who have made Cambridge feel like home (and many of whom I've had the pleasure of sharing a home with here), thank you for being you. There are too many people who have shared their friendship and kindness with me for me to list them all, but I am so grateful for each and every one of you.

Finally, to my wonderful partner, Jason. You inspire me every day with your warmth, compassion, humility, and wisdom. I can feel myself becoming a better person just for being with you, and I cannot imagine navigating this final stretch of my PhD without you. I love you so much and cannot wait to see what the next chapter has in store.

I also owe a great many thanks to all of the non-humans who enrich my life. Thank you first to the monkeys on Cayo Santiago. To the bonobos, Kanzi, Elikya, Nyota, Maisha, Mali, Clara, and Teco, thank you for sharing your time, your patience, and your senses of humor with me. It has truly been a pleasure. Last of all, I am tremendously grateful for my amazing dog, Ziggy. Since I found you in Puerto Rico five years ago, you have been with me every step of the way. You have attended more lab meetings, lectures, and departmental events than any dog I know, and your patience is boundless. It is one of my greatest pleasures to be your human.

Chapter 1 Introduction

Organisms are born into a world of dense information. In order to learn, they must determine what information to attend to and what to ignore. Witnessing something that violates one's expectations can serve as a good indication that there is something to be learned (Busch & Legare, 2019; L. Schulz, 2015; Stahl, 2015; Stahl & Feigenson, 2015, 2017, 2019). Specifically, seeing something surprising can trigger a search for explanation, and a young learner can take advantage of these situations to scaffold their understanding of the world around them.

Newborns do not start out as a completely blank slate; they have some innate knowledge built into the system that can serve as a foundation for subsequent learning (Spelke et al., 1992; Spelke & Kinzler, 2007). This knowledge, termed "core knowledge", consists of domain-specific systems that impact organisms' expectations about objects, geometry/navigation, agents, and number (Spelke, 2000; Spelke & Kinzler, 2007). These systems seem to be innately endowed, as they arise early in development and require minimal environmental input. For example, newborn chicks in controlled rearing conditions seem to have expectations that are consistent with core knowledge of objects (Chiandetti & Vallortigara, 2010; Regolin & Vallortigara, 1995). Core knowledge systems also seem to be evolutionarily ancient, as they have been observed in a wide variety of non-human animal species, from fish to apes (eg: Hauser & Carey, 2003; Regolin & Vallortigara, 1995; Santos & Hood, 2009; Santos, 2004; Spelke & Lee, 2012), and they are present across human cultures (Dehaene et al., 2006; Everett, 2005; P. Gordon, 2004). This rich starting state provides a basic model upon which experience can build through the course of development and can help young learners to gain traction in identifying what to attend to for subsequent processing and learning (Spelke et al., 1992; Stahl, 2015).

Witnessing something that is inconsistent with what one expects, including expectations from core knowledge, could help to address the problem of knowing what to prioritize to learn in order to gain a more nuanced understanding of the dynamics of objects in the world around them and how they interact. In support of this theory, evidence has revealed that children and infants (1) look longer at scenes that include a violation of expectations than those that do not (Wang et al., 2004 for review), (2) engage in exploration and information seeking following such a violation (Baldwin & Markman, 1993; Bonawitz et al., 2012; Busch & Legare, 2019; Legare, 2012; Legare et al., 2010, 2016; Legare & Gelman, 2014; Pieraut-Le Bonniec, 1985; Schulz, 2012; Schulz et al., 2008; Schulz & Bonawitz, 2007; Sim & Xu, 2017; Stahl & Feigenson, 2015, 2017; Subbotsky, 2010), (3) tailor their exploration to provide themselves with useful information (Cook et al., 2011; Legare, 2012; Schulz & Bonawitz, 2007; Stahl & Feigenson, 2015), and (4) show signs of increased learning following surprising events (Stahl & Feigenson, 2015, 2017). These findings are all consistent with the proposal that expectancy violations can help children prioritize certain information in their environment for enhanced learning.

Specifically, this mechanism can aid in the development of flexible causal reasoning, or the ability to think about the effect that one event can have on another (Bonawitz et al., 2012; Buchsbaum et al., 2012; Legare, 2012; Stahl & Feigenson, 2015). For example, based on the core knowledge principle of support, an infant would be surprised to see a ball hovering in midair. This surprise might then motivate them to explore, and this exploration might reveal that there is a magnet on the back of the ball that holds it to a magnet on the wall. A young learner might then gain insight into magnets might cause objects to move, and this can be integrated into their mental representations about objects and how they interact with their environment (Figure 1.1).

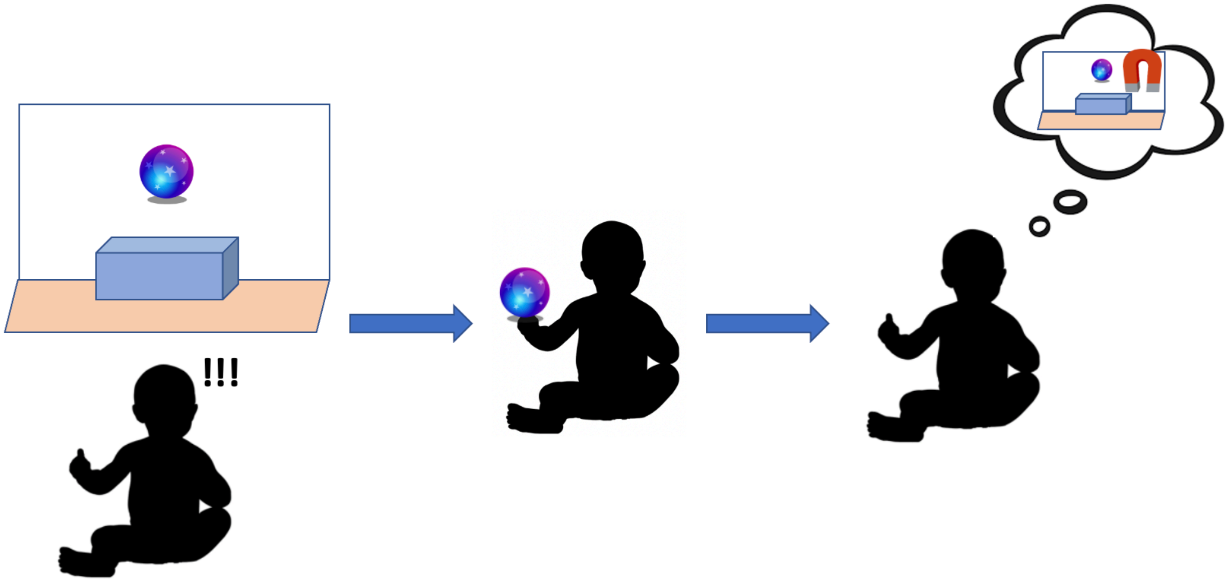

Figure 1.1 Graphical representation of how surprise-induced exploration might inform causal reasoning. An infant might see an object that appears to be floating, explore that object in search of an explanation for the surprising event, and gain new information that helps make sense of the event.

This theory leaves a sort of puzzle, however. Like humans, non-human primates (from here on, primates) have expectations, including those based in core knowledge, and have demonstrated evidence of surprise to violations of these expectations (eg: Santos, 2004; Santos & Hauser, 2002; Spelke, 2000; Uller et al., 2001). For example, like human infants, monkeys have been observed to look longer when an object appears to roll through a wall, violating the core knowledge principle of solidity, than when it appears to stop when it comes into contact with a wall (Santos, 2004). Thus, there is evidence of a similar starting point, as both humans and primates share basic expectations of how objects interact in the world. Humans, however, go on to develop far more complex causal reasoning (Buchsbaum et al., 2012; Penn & Povinelli, 2007; A. Seed et al., 2011; Visalberghi & Tomasello, 1998). Following the argument that witnessing a violation of expectations provides an opportunity to scaffold causal reasoning, there may therefore be something distinct that occurs when a human learner is faced with such a surprising event. The question then remains of whether and how humans might be processing this information differently once they recognize a violation that allows them to scaffold this knowledge into causal reasoning.

This question, how humans and non-humans might be learning differently from surprise (violation of expectations), is the main question of my research program. In this thesis, I argue that humans and primates share our innate foundations for reasoning about the physical world, but humans may be unique in their propensity to seek explanations when their expectations about the physical world are violated. We aimed to design studies that could be conducted with both humans and non-humans to allow for direct comparison, when possible. In Chapter 2, we sought to validate the canonical measure of violation of expectations, looking time, with both human infants tested online as well as with rhesus macaques in a semi-free-ranging habitat. In Chapter 3, we used an innovative paradigm to demonstrate that bonobos can use core knowledge to reason about their world, and that they therefore have expectations about how objects interact in the world, which could then be violated. And in Chapter 4, we discuss evidence that bonobos unlike human infants, do not prioritize objects that violate their expectations for subsequent exploration.

1.1 Surprise: when expectations are violated

Surprise results when an individual witnesses or experiences something that is inconsistent with their expectation(s) of that event (Barto et al., 2013; Ekman & Davidson, 1994). Experiencing surprise does not depend on any previous experience with the situation, as with witnessing violations of core knowledge. Surprise can be viewed as consisting of two elements; the first is the computation of the discrepancy between what one expected and what one observed, and the second is the subjective experience that accompanies that mismatch. Throughout this thesis, we will primarily be dealing with the first of these two components as it is more operationalizable and objective and has been previously identified in primates. The subjective experience of surprise is much more ambiguous and previous research has not decisively determined whether primates experience this feeling, as there are no known patterns of facial expressions or physiological changes associated with the experience of the feeling of surprise[1] (Kret et al., 2020). The focus on surprise here is that surprise may actually serve an adaptive function, in that it can help direct a young learners' attention towards relevant learning opportunities and can inspire a search for a causal explanation (Charlesworth, 1964; Darwin, 1872; Tomkins, 1962).

1.2 Surprise as a wedge into learning

Surprising events have been shown to increase learning, particularly during development. Witnessing a violation of expectations of various types of events has shown to boost learning in both children and in infants and has been observed in more controlled and more naturalistic settings (see Stahl & Feigenson, 2017 for review). For example, human children more effectively learned a novel word (a verb or noun) when it described an object that participated in a surprising event rather than an expected one, and learning was targeted in that it did not extend to irrelevant objects (Stahl, 2015; Stahl & Feigenson, 2017). A follow-up study determined that this increased word-learning effect held only when the surprising event was impossible but not improbable. Children were presented with a gumball machine apparatus that varied in the proportion of colored gumballs inside. In one condition there were 50% purple gumballs and

---

[1] Here we specifically refer to surprise as it relates to violation of expectations, not startling.

50% pink gumballs, in another it was 100% of one color, and in a third it was 10% one color and 90% the other. Only children who witnessed an impossible event, for example seeing a pink ball come out of a machine with only purple balls, learned the new word for the pink ball. The children in the improbable condition, where a gumball of the color that made up 10% of the population came out, showed no increase in learning over the children in the 50/50 condition (Stahl & Feigenson, 2019). While word learning itself is not necessarily relevant for causal reasoning, this study demonstrates something important about the link between surprise-induced learning and causal reasoning: impossible events, but not improbable ones, evoke the question, *how* did that happen? When the low probability outcome emerges, it might seem unlikely, but one would not need to explore to figure out how that possibly occurred. However, when the gumball emerges that was not part of the initial population at all, this may trigger a search for explanation that then heightens learning.

There is evidence that infants, even within the first year of life, also use surprise as a cue that there is something to learn. Stahl & Feigenson (2015) demonstrated that witnessing an event that is inconsistent with principles of physics increased infants' learning of new information, and that this learning was specific to the object that was involved in the violation. Specifically, they showed that infants were better able to map a hidden auditory property onto an object when that object had previously behaved in a way that violated expectations.

There are several different mechanisms by which surprise might serve to boost subsequent learning, and these are not mutually exclusive. For one, it could be the case that seeing something surprising merely serves to increase attention to the event (Fazio & Marsh, 2009) and attending to the event more provides more of an opportunity to learn from it (Reisenzein et al., 2019; Stahl, 2015; Turk-Browne et al., 2005). That is to say that shifting attention away from other stimuli and towards the surprising event, which in fact have been shown to be two separate processes (Horstmann, 2006), frees cognitive resources so that they are available to attend to and analyze the unexpected event. In other words, enhanced learning is merely a byproduct of increased attention, as surprise leads to increased attention which leads to learning, with no direct link between surprise and learning at all. A most basic piece of evidence that surprising events elicit additional attention is that individuals tend to look longer at surprising events than unsurprising events (eg: adults: Horstmann et al., 2016; Horstmann & Herwig, 2015; Retell et al., 2015; infants: Baillargeon, 1995; Spelke et al., 1992 non-humans: Santos & Hauser, 2002). Additional evidence also comes from findings that witnessing a surprising event decreases people's performance on parallel tasks, indicating that they really are shifting attention away from other input to allocate it to the surprising event (Horstmann, 2006). Allocating attention to an event that violates one's expectations can also then serve to increase memory of that event in the future (Fazio & Marsh, 2009; Greve et al., 2017; Munnich et al., 2007), illustrating the extent to which these are meaningful and lasting opportunities for learning.

A second possibility is that individuals might not only *attend* more to unexpected events, but they might actually *process* unexpected events more deeply and therefore encode more information about the object that participated in the event (Reisenzein et al., 2019; Stahl, 2015). Experiencing something unexpected could increase uncertainty of one's environment, which might in turn increase the cognitive effort required to assess and analyze the scene, thus leading to deeper encoding and memory for the event (Csink et al., 2021).

Reisenzein et al. (2019) propose that it is the cognitive experience of surprise that directs the attentional shift towards the surprising event, and that the subjective experience of surprise elicits curiosity and therefore incites a deeper processing of the event and a search for an explanation. If it is indeed the case that primates do not have the same subjective feeling of surprise as us, this could potentially explain a difference in their tendency to learn from surprising events.

1.2.1 Surprise as a wedge into learning in non-humans: Prediction error

The idea that surprise facilitates learning is not a new one; it is a principle that has been explored in the context of prediction error in conditioning studies with non-human animals for decades. Models of prediction error have shown that a mismatch between what an individual expects and the outcome increases the rate and efficacy of learning and contributes to cognitive development (McClelland, 2002; Pearce & Hall, 1980; Rescorla & Wagner, 1972). That is to say that animals better learn the association between the conditioned stimulus and the unconditioned stimulus when the conditioned stimulus does not always accurately predict the consequences. Rather, learning is more effective when the consequences are not always predictable, as prediction error can serve to increase animals' attention to the cues that are present during the learning trial.

While it is currently unknown whether this phenomenon operates by the same mechanism as does surprise-induced learning in humans (Stahl & Feigenson, 2019), it does seem like they are based on the same principle. What is unique about the perspective that human infants show enhanced learning following a violation of expectations is (1) that young learners might take advantage of innate or naturally acquired expectations, as opposed to learned ones, and (2) that they strategically harness this surprise to guide their own exploration and learning in search of an explanation, rather than merely passively learning more due to attentional shifts.

1.2.2 Bayesian learning: A theoretical framework for surprise as a wedge into learning

Bayesian learning provides a theoretical approach to explain the mechanism of surprise-induced learning, as it addresses how organisms use evidence to update their beliefs and understanding about the world (Perfors et al., 2011). According to this theory, an organism initially has some *prior* and then integrates new observations and information to update these beliefs into a *posterior*, using Bayes' theorem (note that this is not proposed to be a conscious or explicit process) (Barto et al., 2013; Itti & Baldi, 2009). Reasoning by Bayesian principles would serve to guide the learning process in a handful of ways. First, it helps to narrow the hypothesis space, as priors can help organisms to limit the range of possibilities that they consider in order to understand an event (Sobel et al., 2004). Second, it promotes certain instances over others for learning, specifically those where an observation is inconsistent with the individual's priors, resulting in Bayesian surprise (Itti & Baldi, 2009). This surprise serves to increase attention to the event (Itti & Baldi, 2009), and indicates that new learning is needed (Courville et al., 2006). Finally, it is proposed that Bayesian surprise increases curiosity and exploration (Bonawitz et al., 2012; Croker & Buchanan, 2011; Schulz et al., 2007, 2008).

1.3 Surprise-induced exploration

Individuals do not just passively learn more following a surprising event, they might also engage in more active information-seeking following a violation of expectations. The role of surprise in exploration has been studied most extensively with preschool and elementary school children. Children are motivated to pay special attention to observations that are inconsistent with their prior knowledge and expectations, and this leads not only to increased looking time to these events, but also increased exploration and even targeted experiments that seek to disambiguate what they witnessed, the results of which they are able to integrate into their understanding (Bonawitz et al., 2012; Busch & Legare, 2019; Legare et al., 2010).

Chandler & Lalonde (1994) conducted one study that addressed this in the context of a violation of core knowledge. They showed 3-5-year-old children an object that was then placed behind a wall. The wall then fell all the way back, appearing to violate object permanence and solidity. The majority of children who witnessed this event searched behind the wall for clues as to what had happened, when given the opportunity. The children's search for an explanation for this surprising event demonstrates that they are not necessarily restructuring their conception of how the world works based on this instance of "magic", but rather may be searching for evidence that helps to make sense of the event within the framework of their prior conception of the world.

Subsequent studies demonstrated that children's expectations need not be innate in order for prior beliefs to have such an effect. Bonawitz et al. (2012) conducted a study with 4-7-year-olds to see how their theories of object mass would impact their exploration. Specifically, they compared children who believed that an object's balance point is always at its center (typically younger children) to children who correctly believed that an object's balance point depends on where the center of mass is (typically older children). They then showed the children one event that was consistent with their theory and one that was inconsistent. They found that children preferentially explored the object that had surprised them, consistent with their own prior beliefs. In other words, children who believed the balance point was at the center preferred to explore the object that balanced at one end, and children who believed the balance point was at the center of mass preferred to explore an object that balanced at the center despite an uneven weight distribution.

Other studies extend this finding by showing not only that children prefer to explore an object that surprised them, but that they do so in a way that isolates variables to update their knowledge and comprehension. One study utilized children's expectations about shadows, differentiating Rule 1 children, who thought that only the size of the object was important in determining the size of the shadow, from Rule 2 children, who believed that both the size of the object and the distance from the light source were important. After determining which group a child fell into, they would provide them with one example that was either consistent or inconsistent with their belief. The children then had the opportunity to play with the objects and the light box. They found that only the children who had seen an inconsistent outcome attempted to isolate variables, changing only the size of the object but not the distance, and vice versa (van Schijndel et al., 2015). Another example of this phenomenon comes from Legare's (2012) study, which used a common paradigm called the "blicket detector" task. In this study, they showed 2-6-year-olds a box that lights only when one type of object is placed on top (a "blicket") but does not light up for any other objects. They then showed some of the children either a blicket or a non-blicket that caused the box to light up and then allowed the children to explore the objects and the box. They found that the children who had had their expectations violated by seeing a non-blicket light up the box seemed to test what had happened by exploring the box and objects in a more targeted way than the children who had not had their expectations violated. Together, these studies demonstrate that young children are motivated to explore when evidence is surprising and that they do so in a targeted way that helps them to better understand what they had just witnessed so that they can update their model of the world.

While far fewer studies have sought to address surprise-induced exploration and hypothesis-testing emerge earlier in life, a handful of studies have indicated that even infants engage in these information-seeking behaviors. At the most basic level, it has been proposed that infants' propensity to look longer at surprising events reveals a tendency to attend more to, and to seek additional information from, instances that violate their expectations (Stahl, 2015). In a more direct test of this phenomenon, Pieraut-Le Bonniec (1985) found that by 9 months of age, infants preferentially explored an object that appeared to be concave but was actually altered so that it was not, as opposed to an object that actually was concave. Furthermore, it appeared that the infants were engaging with the object to better understand the disparity between what they saw and what they felt. Sim & Xu (2017) used a crawling paradigm to examine selective exploration in 13-month-olds. They showed babies two boxes of balls of different colors and either drew out four different colored balls or four of the same color, which looking times indicated the infants found surprising. They then allowed the babies to crawl towards the boxes, and they found that the majority of babies crawled toward the box that had surprised them.

Stahl & Feigenson (2015) sought to assess exploration and hypothesis-testing in 11-month-old infants following a violation of expectations. They investigated whether infants acted to gain information on their own by selectively exploring violating objects and whether they did so in ways that were specific to the violation they had witnessed. They found support for both of these hypotheses. Infants who witnessed a violation of expectations spent more time engaging with the violating object relative to a novel object, whereas infants who had not seen such a violation spent equal time engaging with each type of object. Furthermore, infants who saw an object appear to pass through a solid wall were more likely to bang the target object and those who saw an object appear to roll off a surface and hover unsupported were more likely to deliberately drop that object on the table or floor. Thus, infants tailored their exploratory behavior to the type of violation they had seen. Follow-up studies determined that these effects were specific to the object that had behaved in a surprising way, as they did not generalize to objects of the same kind (Stahl & Feigenson, 2019). This is important, because in generating a causal model of an event, it is important that it applies only to that specific event. For example, if an infant is surprised by a ball that appeared to roll through a wall, further investigation might reveal that there was actually a trap door that allowed the ball through. This provides an explanation of what caused the ball to continue rolling in this specific circumstance, but it would be incorrect to assume that every apparent violation of solidity is also caused by a trap door.

In addition to seeking direct evidence for the violation they just witnessed, infants might also seek addition information from social partners. Walden et al. (2007) found that infants engage in more social referencing during surprising event. In other words, infants were more likely to look at their caregiver when witnessing an object that behaves in a way that violates their expectations as compared to when faced with one that does not.

To my knowledge, only one study has been published that addresses selective exploration and hypothesis testing in primates. In this study, children and chimpanzees were taught to stand blocks up on one end to earn a reward. At test, they were presented with a number of blocks to try to stand up, but one of them had been subtly altered so that this was impossible. They found that both children and chimpanzees spent more time with the sham block, which is not surprising since it was more difficult. However, they also found that both groups were also more likely to try moving to a different location with the sham block than with the others and that they engaged in more close inspection of the sham block. One difference that they identified was that the apes were not as consistent in their increased exploration across studies, and that children were more likely to look closely and feel for a difference between the blocks (Povinelli & Dunphy-Lelii, 2001).

1.4 Causal learning and reasoning

1.4.1 The link between surprise-induced learning and exploration and causal learning

Surprise-induced learning and exploration are proposed to help learners to identify underlying causal structures of events (Schulz & Bonawitz, 2007) which then serves to develop causal reasoning more broadly.

Before proceeding any further, though, it is important to define exactly what is meant by the different terminology that is used in reference to causality. One difficulty with this is that different terms are used by different scientists, and it is often unclear whether they are referring to the same phenomena. By most accounts, however, *causal knowledge* and *understanding* involve knowing that two objects or events are somehow related to one another, as well as recognizing that there is some force or structure that defines the relation between the two (Call, 2010; Call & Tomasello, 2005; Visalberghi & Tomasello, 1998). *Causal learning* is the process of learning these relations (i.e. their causal structure) and in coming to understand that this structure and the underlying forces that lead to the event can be used to make predictions or interventions (Bullock et al., 1982; Visalberghi & Tomasello, 1998). Finally, *causal reasoning and causal inference* refers to the ability to discern underlying causal mechanisms and make predictions about future outcomes or counterfactuals upon first observation, without any additional information or exploration (Call, 2010).

Evidence from young children support the proposal that surprise-induced exploration might serve to promote learning about causality. The evidence that children reveal in their targeted exploration and interventions, such as in the previously-described studies, does provide information about causality (Bonawitz et al., 2010; McCormack et al., 2016; Schulz et al., 2007).

Children are also able to make judgments about the strength of a causal relationship based on their interventions and tend to trust the results of their experimentation over the baseline probabilities (Kushnir & Gopnik, 2005). When children are asked to verbally explain an event they just witnessed, they provide more detailed explanations for surprising, as opposed to unsurprising events, and are more likely to make reference to underlying causal mechanisms (Legare et al., 2010). The fact that impossible events have an even stronger effect on exploration and learning than do merely unlikely or counter-intuitive events (Stahl & Feigenson, 2019; Subbotsky, 2010), supports this idea, as well. Seeing something that is truly impossible increases the need for a causal explanation to understand how that event possibly occurred. Seeing an improbable event may be surprising, but one may not need to search for an underlying cause to explain how it could have happened at all (Stahl & Feigenson, 2019). In addition, the evidence that children and infants do not generalize the new information they learn about a surprising object to other objects of that kind (Stahl & Feigenson, 2019) indicates that there is some recognition that there was some specific cause in that instance that may not apply to all instances involving objects of that kind. For example, a ball disappearing from one location and reappearing in another might be explained by someone moving that object, but that cause and effect would not necessarily apply to all balls.

Gaining insight into the causal structure of an event does not necessarily mean that an individual will have to update their beliefs. In fact, discovering hidden causal variables might actually reinforce prior beliefs, if it serves to explain away what seems to have occurred (Bonawitz et al., 2010). For example, if a child saw an object go behind a wall, but the wall then falls all the way back as if nothing were behind it (as in Chandler & Lalonde, 1994), this does not necessarily mean the child needs to update their beliefs about object permanence or solidity.

Exploration might reveal that the object fell through a small trap door on the table, which then caused the wall to fall all the way back when it was released.

The idea that the increased exploration and hypothesis testing triggered by surprise helps to scaffold causal learning fits into the Bayesian framework (Gopnik & Tenenbaum, 2007; Tenenbaum et al., 2006). This framework combines features of bottom-up approaches, which emphasize the importance of statistical cues of causality independent of prior knowledge with top-down approaches, which emphasize the extent to which prior knowledge of causal mechanisms guide causal inference and help to explain how individuals can make correct causal attributions from limited examples (Tenenbaum et al., 2006). According to a causal Bayesian network, after witnessing an event that conflicted with their priors, learners would test hypotheses of potential causal structures by isolating potential variables and observing their effects and interactions (Gopnik & Tenenbaum, 2007).

For this thesis, which investigates the potential mechanisms that underpin the ability to translate a violation of expectations into causal reasoning by comparing humans and primates, it is important to have an understanding of the presence of causal understanding and reasoning in children and apes.

1.4.2 Causal learning and reasoning in humans and apes

Within the first few years of life, children show evidence of being able to think about causality. They are able to explain events using causal language, understand others' claims about causality, predict future events based on the relationships between objects and events (Bonawitz et al., 2010; Gopnik et al., 2004 for review), and they can make causal judgments by using the events they witness to update their beliefs (Kushnir & Gopnik, 2005; Schulz et al., 2007; Schulz

& Gopnik, 2004; Sobel & Munro, 2009). In fact, from a very young age, it seems that children reason about causality according to the same principles as adults: (1) *determinism*, that events must have a cause, (2) *priority*, that one event or object causes the other, and not necessarily vice versa, and (3) *mechanism*, that there is some particular underlying cause that can be identified (Bullock et al., 1982).

There is even some evidence of the early foundations of causal understanding in infants. Human infants around seven months of age were found to be sensitive to the structure of a cause-and-effect relationship. They were habituated to a launching event, where Object A hits into Object B, causing Object B to move, and they were then shown the reverse event, such that Object B appears to make Object A move. Infants looked longer when the direction, and therefore the causality was switched, but they did not look longer if there was a delay between when the objects made contact and when the second object began to move, such that the event did not appear to be causal (Leslie & Keeble, 1987). In a study with older babies, 16-month-olds were shown that when a button is pressed on a toy it plays music. They were then handed an object that looked identical to the original one or a new object that differed only in color. When the infant then tried to press the button, no music played. Infants who had an identical object seemed to reason that it was the agent that was the problem, as they attempted to hand the toy to their parent. Those who had a slightly different toy requested another toy (by pointing or pulling it closer), indicating that they believed the toy to be at fault and not themselves. The authors take this as evidence that the infants were making causal attributions about the failure of the toy based on what little evidence they had (Gweon & Schulz, 2011).

Apes have also succeeded at tasks that require at least some degree of causal understanding. In an experiment modeled after Leslie & Keeble's (1987) study with human infants, apes were shown causal sequences that were either “natural” or “unnatural”, such as a hand pushing a banana versus a banana pushing a hand. They found that when apes were habituated to the natural event and then saw the unnatural event, they looked for far longer than when they were habituated to the unnatural event and then saw the natural event. This demonstrates that they were not merely looking longer because of the novelty of the reversed event, but they were in fact processing the causation of the event (O’Connell & Dunbar, 2005). Premack & Premack (1994) had apes watch as two cups that were equidistant were baited with two different foods. The apes’ view of the cups was then blocked and when it was revealed again, an experimenter was eating one of the two foods. The apes were then free to explore the cups for food, and they found that the majority of individuals tended to check the cup that would have contained the other food. They concluded that the chimpanzees were therefore able to complete the causal representation based on the causal sequence (note: others disagree with this interpretation).

A number of studies have used the “trap-tube task” where there is a tube with a piece of food inside, but there is a hole in the bottom of the tube that the food item will fall into and be lost if it passes over. Participants must use a tool to push the food item out, but they should push it from whichever direction will not lead it to pass over the trap. While the majority of early studies with this task found that primates could not pass using causal reasoning (eg: Visalberghi & Limongelli, 1994), subsequent studies reflected greater success, particularly when they removed potential task constraints, such as the requirement of tool use or of using a pushing rather than pulling motion, and found that primates could succeed (Girndt et al., 2008; Limongelli et al., 1995; Mulcahy & Call, 2006; Seed et al., 2009). Another common paradigm is the “stick-and-hook problem”, where participants must pull a stick to retrieve an object with which it does not have direct contact. A variety of primate species, including great apes, have succeeded at this task, choosing to pull the stick where the reward is within the hook of the stick over one where the stick does not have a hook, or the object is outside of the hook(Call & Tomasello, 2005; see Tomasello et al., 1997 for review).

Still further evidence of causal understanding in apes comes from a study in which chimpanzees were shown a series of steps that could be performed to retrieve a food reward from a puzzle box. They found that the chimpanzees copied every step when the box was opaque and they could not see the effects of their actions, but when the box was transparent and they observed that their actions did not cause any change at all, they omitted those steps and went directly for the food reward (Horner & Whiten, 2005).

A final piece of evidence comes from a study in which apes were presented with a small balance beam that had a cup on each end. In one setup, the balance beam was functional, such that the heavier cup would move downward, and in a second setup the beam was fixed, such that one side was always lower than the other. They found that only in the first of these two conditions, chimpanzees consistently chose the lower cup, indicating that they reasoned that it was the weight of the food within the cup that would have made the mobile beam move in that direction (Hanus & Call, 2008).

Despite these successes, human causal reasoning is often supposed to be unique, though there is little consensus in the literature about what specific features of causal reasoning set humans apart. There are some common themes that emerge, however. Human causal reasoning is proposed to be superior in that humans (1) apply causal reasoning to the infinite range of scenarios that they encounter (Buchsbaum et al., 2012), whereas the majority of primates' successes on causal reasoning tasks have been found in the context of foraging or food retrieval and dominance (Seed et al., 2011), (2) use causal Bayes nets and Bayesian learning of causal models (Buchsbaum et al., 2012; Penn & Povinelli, 2007), (3) are more sophisticated in their use of causal regularities that they observe (Penn & Povinelli, 2007), (4) have causal theories that are more abstract and theory-like whereas non-humans tend to rely on relations that can be observed (Penn & Povinelli, 2007), (5) can use arbitrary cues of causality, rather than functionally-relevant cues (Seed et al., 2011), and (6) are able to draw analogies between scenarios with the same underlying causal structure (Martin-Ordas et al., 2008; Penn & Povinelli, 2007). It has further been suggested that non-humans lack the necessary underlying abilities for flexible and abstract causal reasoning, namely counterfactual thinking and mental time travel, and that these abilities developed in tandem with the extension of the juvenile period throughout human evolution and the opportunity for pretend play that this allows (Buchsbaum et al., 2012). While the exact ways in which human causal reasoning varies from non-humans are still highly contested and remain somewhat unclear, it does seem to be the case that humans generally outperform even their closest relatives.

While apes are able to use abstract structural knowledge of object properties within a causal framework, such as predicting the effects that objects will have on one another in order to find food (Seed et al., 2011), it remains the case that causal reasoning appears to be more complex and prolific in humans than in non-human animals, being used more flexibly and across a wider range of situations. Specifically, non-humans seem to struggle to infer the underlying causal mechanisms of events, even if they understand that a given cause leads to a given effect. They also cannot transfer causal knowledge across different scenarios or understand that analogous scenarios share an underlying causal mechanism (Martin-Ordas et al., 2008; Martin-Ordas & Call, 2009; Vaesen, 2012). In other words, it seems that primates have a harder time reasoning and the causes that they cannot see and must think abstractly about. Primates do seem to recognize impossible events as surprising, as they have shown increased looking to scenes that violate their expectations, so in theory they too could use these scenarios as a cue for learning.

1.5 Preview of Chapters 2-4

The goal of this thesis is to gain traction into the question of how humans and primates might come to understand how their world works and how they might act on it in ways that provide themselves with additional information to scaffold their understanding. All studies presented as part of this thesis were designed to be run nearly identically in both human infants and primates. All tasks are non-verbal, require no prior training or experience, and suitable for infants, monkeys, and apes.

Chapter 2 Validating looking time as a measure of expectancy violation in human infants and primates

2.1 Background

2.1.1 Core knowledge: innate expectations about the world

Core knowledge consists of at least four distinct systems, each dedicated to representing (1) objects and how they interact with the world, (2) agents and the utility of their actions, (3) numbers and their relations to one another, and (3) geometry in a spatial layout (Spelke & Kinzler, 2007). The focus here will be on infants' and primates' understanding of and expectations about objects. Infants from a very young age are able to reason about objects as complete and individual entities, even when they are out of view, and they expect them to adhere to a number of principles of physics (Lea et al., 1996; Regolin & Vallortigara, 1995; Spelke, 2000; Spelke et al., 1992; Valenza et al., 2006). These include *solidity*, that one solid object cannot pass through or occupy the same space as another solid object, *support*, that an object cannot hover in midair without falling down, *continuity,* that objects move in a straight manner without spontaneously changing direction or disappearing in one location and appearing in another, and *inertia*, that an inanimate object cannot spontaneously start moving on its own (Spelke et al., 1992; Spelke & Kinzler, 2007). Infants track and reason about objects from a very young age and are able to represent objects and their properties even when they do not have continuous visual access to them.

Much of what we have come to know about infants' understanding of objects and how they behave have come from studies that measure how long infants look at different events and outcomes. Typically, studies that seek to assess infants' understanding of object utilizes the violation of expectation (VOE) method. In this paradigm, infants are habituated to a possible event, such as an object being pushed partway across a box. Often infants are shown the event repeatedly until they reach a certain threshold of decreased visual attention to the display, indicating that they have habituated to the scene. Participants are subsequently shown events that are either possible (eg: an object is pushed from one edge of a box to the other) or impossible (eg: an object is pushed beyond the surface of a box and remains suspended in midair). If infants recognize the impossible event as a deviation from what they expected, they should dishabituate and therefore look longer at the event. These paradigms are typically designed such that the impossible event is more perceptually similar to the habituation events than is the possible event so that longer looking cannot be attributed to perceptual novelty. Studies have also demonstrated that the VOE method works without any habituation phase (Wang et al., 2004), which provides evidence that this method is actually allowing access to infants' prior expectations and not just what they have learned through the course of the study. Hundreds of studies have used this method to uncover pre-verbal infants underlying expectations, including those from core knowledge of objects (see Spelke & Kinzler, 2007 for review; see Stahl, 2015 for reviews).

This method has also been successfully used with primates and has demonstrated that they too have expectations about objects based in core knowledge (eg: Cacchione & Krist, 2004; Murai et al., 2011; Santos & Hauser, 2002). Because it is often difficult to monopolize a primate's attention for studies in the settings in which they are typically tested, VOE studies with primates often utilize a briefer familiarization portion of the experiment, rather than seeking to achieve a certain habituation criterion. This method has been successful in uncovering primates' underlying expectations of objects just the same, such as that they expect an object to stop when it hits a wall rather than to be able to pass through a wall (Santos & Hauser, 2002).

Looking time is a powerful tool in that it is a sensitive measure that can be used flexibly without requiring any language or training. Looking is typically automatic and is therefore less susceptible to bias or task demands from the participants. This makes it a very useful tool for gaining an understanding of both developmental and cross-species comparisons that aim to shed light on the foundational understanding of objects and how they act in the world.

2.1.2 Core knowledge of object solidity

Infants as young as 2.5 months old have demonstrated expectations of object solidity. In an early study of infants' core knowledge of objects, they were shown a scene where an object rolled down a ramp and behind a barrier. When that barrier was lowered, infants looked longer when the object was on the opposite side of a wall that blocked the ramp, implying that the object had rolled through the wall, than when the object was at rest at the near side of the wall (Spelke et al., 1992). It seems, however, that infants' understanding of object solidity does not come online all at once, as infants do not show the violation of expectation effect with vertically falling objects until 4 months of age. That is to say that even once infants look longer when an objects appears to pass through a barrier wall when traveling horizontally, they do not yet look longer when an object is dropped behind an occluder and appears to pass through a shelf to end up underneath (Spelke et al., 1992).

Another common paradigm for assessing expectations of object solidity entails habituating infants to a wall falling forwards and backwards repeatedly. An object is then placed behind the wall and the wall either falls all the way back again, which is perceptually similar to what they saw in habituation but violates the principle of solidity, or the wall stops where it would have come in contact with the object, which is perceptually novel but consistent with the principle of solidity. Infants as young as 3.5 months of age have been shown to look longer at the impossible case over the possible case (Baillargeon, 1987), providing further evidence that they are able to reason about an object's solidity even when it is fully occluded.

Primates also show evidence of expectations of object solidity using a looking time paradigm. Santos & Hauser (2002) tested adult rhesus macaques (*Macaca mulatta*) on two paradigms that were nearly identical to that of Spelke et al. (1992). In one set-up, they showed monkeys a plum rolling down a ramp and revealed it to be either at the near or far side of wall that blocked the track, and in a second set-up they showed monkeys an apple slice being dropped and revealed it either resting on a shelf that would have blocked the apple's trajectory or beneath that shelf. The biggest deviation from Spelke et al.'s original paradigm was that in place of a series of habituation trials, macaques in each experiment were shown two familiarization trials, one with the object in each of the two potential end locations, so that neither was more novel than the other at test. Overall, monkeys tended to look longer when the object was revealed at the far side of the wall blocking the ramp, appearing to have rolled through the blocking wall, and when it was revealed beneath the vertical shelf, appearing to have passed through the shelf when dropped. This provides evidence that adult macaques had expectations about the object's movement that were grounded in the core knowledge principle of solidity.

Interestingly, while looking time measures reveal expectations and understanding of solidity very early in human infancy as well as in adult primates, reaching and exploration paradigms are much less consistent. 2-year-old children make errors in search tasks that depend on reasoning about solidity, both when the object moves horizontally (Berthier et al., 2000; Hood et al., 2000) and vertically (Hood et al., 2000). Adult rhesus macaques were tested on an exploration version of the study described above, where monkeys were shown both horizontal and vertical set-ups designed to probe understanding of solidity. Macaques have demonstrated mixed success on the horizontal version of the task. In one study they were found to search the near side of the barrier wall (Hauser, 2001), indicating that they recognized that the wall would block the object's trajectory. However, in a second study, macaques seemed to fail in the horizontal search task, tending to search wherever they had last seen the object. They did succeed in searching in the correct location when only the two potential locations were occluded but the rest of the track was visible, indicating that they required additional spatiotemporal information about the object and struggled to reason about the motion of the object when it was fully occluded (Santos, 2004). Macaques also tended to fail in the vertical version of the task, searching beneath the shelf that would have blocked the object's trajectory (Hauser, 2001; Southgate & Gomez, 2006). While both young children and macaques did not consistently show evidence of using solidity to correctly reason about the location of the hidden object, in the vertical search task children did not consistently choose one location over the other (Hood et al., 2000), while macaques tended to choose the bottom location (Hauser, 2001).

Cacchione et al. (2009) conducted both the horizontal and vertical version of this task with all four species of nonhuman great ape: orangutans (*Pongo pygmaeus*), gorillas (*Gorilla gorilla*), chimpanzees (*Pan troglodytes*), and bonobos (*Pan paniscus*). They had the apes point to the location where they wished to search, and the outcome was revealed. All four species performed above chance in the horizontal version of the task, more often choosing the location where the object would have stopped before passing through, but they performed at chance in the vertical version of the task, choosing equally the location above and below the shelf.

Overall, it seems that there is compelling evidence that both infants and adult primates have expectations about object solidity that are revealed under certain conditions, but results are not always consistent. Looking time studies seem to reveal greater sensitivity to the constraints of solidity than do search tasks, and having to integrate information about gravity appears to make search tasks using the vertical paradigm particularly difficult. In addition, early studies of infants' expectations based on solidity used video displays, but this was at a time in which infants did not have such abundant access to screens, where principles of core knowledge often appear to be violated, such as on phones, tablets, and personal computers. Subsequent studies that aimed to surprise babies using violations of solidity used live displays rather than video (eg: Perez & Feigenson, 2022; Stahl & Feigenson, 2015), and the effectiveness of video displays outside of a lab setting have recently been called into question (Smith-Flores et al., 2022). Specifically, 15-month-olds did not look longer when an object was revealed on the far side of the barrier wall in the horizontal version of the solidity paradigm described above. However, it is unclear whether this was because (1) the displays were presented over video, requiring a 3-dimensional representation of 2-dimensional display, (2) infants were not shown the full display with the barrier wall in place before the occluder was positioned to block the right side of the display, (3) 15-month-olds are at an age at VOE is no longer an effective measure of solidity understanding, or (4) stimuli were presented to infants at home rather than in a controlled laboratory setting. Notably, other effects in the same set of studies and numerous other studies (eg: Bochynska & Dillon, 2021; Chuey et al., 2021) were found to be consistent with in-lab findings, so the fourth explanation here seems less likely.

The studies presented in this chapter aim to provide a clearer picture of humans' and primates' expectations of solidity using a similar paradigm. We contribute to the literature on human infants with Study 1 by (1) testing infants at an age that is intermediate to those in earlier studies that found evidence of solidity expectations (eg: Spelke et al., 1992) and more recent studies that did not find such evidence (Smith-Flores et al., 2022), (2) testing a relatively large sample size of infants, (3) utilizing a within-subjects design, (4) testing infants on a 2-dimensional display but providing them with full visual access to the display prior to occlusion, and (5) using familiarization trials rather than a habituation period, which is more similar to the design of surprised-induced learning studies (eg: Stahl & Feigenson, 2015) and studies with primates (eg: Santos & Hauser, 2002), such as in Study 2. We contribute to the literature on primates with Study 2 by (1) providing a developmental perspective, testing monkeys across the lifespan rather than solely adult individuals, as in previous work, and (2) using inedible objects rather than food items, making it more comparable to infant studies.

2.1.3 Core knowledge of object support

Infants as young as 3 months old show evidence of knowledge of support, looking longer when an object is pushed past a supporting surface, such that it appears to float without falling, than when an object is pushed the same distance but remains supported or continues to be held by a hand throughout (Needham & Baillargeon, 1993). The ability to reason about the more nuanced aspects of support appears to unfold throughout development. While 3-month-olds look longer when an object is pushed beyond the surface with which it had contact, Spelke et al. (1992) found that when an object was dropped behind an occluding wall, 3-4-month-old infants looked equally long when the object was then revealed to be floating in midair and when it was revealed to be resting upon a surface. When infants are 4.5-5.5 months of age, it seems they begin to reason about the *type* of contact that an object has with a surface, rather than just discriminating between whether or not it has contact with a surface. Around this age, infants begin to look longer when an object is placed against the side of a supporting platform without support from underneath than when it rests directly on another surface (Baillargeon, 1995). At 6.5 months infants begin to reason about the *amount* of support that is required for an object, looking longer when only 15% of an object is in contact with a platform without falling, as compared to 70% of the object remaining in contact with the platform. Younger infants do not discriminate between these differing degrees of contact, looking longer only when the object has no contact with the platform, but not when it has only minimal support (Baillargeon et al., 1992). A more recent study provided evidence that even older infants continue to show evidence of support understanding in a VOE task, as 15-month-old infants looked longer when an object was pushed past the end of a platform than when it maintained contact with the platform throughout (Smith-Flores et al., 2022).

Primates also seem to use the principle of support to reason about object relations. Research with great apes using a VOE paradigm showed that chimpanzees look longer at a display where an object is unsupported and does not fall than when it remains supported (Cacchione & Krist, 2004; Murai et al., 2011), and also at a display where the object had too little contact with its support (15%) but does not fall compared to when it is mostly in contact with the support (70%). Interestingly, they did not seem to discriminate between different types of support, looking equally long when the object maintained contact with the side of the platform as when it had an additional support beneath it (Cacchione & Krist, 2004). Murai et al. (2011) also found that Japanese macaques (*Macaca fuscata*), like chimpanzees, look longer when an object does not have any contact with a supporting platform, but not when it has support only from the side and not from underneath.

2.2 Study 1: Infants

Study 1 assesses whether infants look longer to violations of solidity and support when tested in their home environment on a 2-dimensional display over Zoom. If infants represent the 2-dimensional scenes as 3-dimensional events and are reasoning based on principles of solidity and support, they should look longer when an object appears to pass through a wall or hover in midair, as compared to when it appears to stop when it comes into contact with a wall or remains supported.

2.2.1 Methods

Participants

Our sample included 61 infants, ages 8.0-10.5 months (33 females, mean age = 9.3 months, age range = 8.2-10.5 months) over Zoom. One additional infant was tested, excluded, and replaced due to inattention and technological issues related to video playback. All infants were recruited from our existing database of families who have expressed interest in participating in studies. All participants must have reached at least 35 weeks gestation in order to participate. Parents received a $5 Amazon gift card after the session to thank them for their participation in the study. This research was reviewed and approved by the Harvard Institutional Review Board (IRB13-2364) and adhered to all guidelines as such, including obtaining informed consent from parents prior to testing.

We conducted power analyses based on the effect sizes from our previous pilot sample to determine a sample size that would give us sufficient power to detect an effect, with an alpha of .05. In this pilot study (n=8), using the same methods described below, we found that infants who saw an unexpected event looked at the display for 63.36% of the allotted time on average (62.29% for solidity, 65.5% for support), whereas infants who saw an expected event looked at the display for 35.86% of the allotted time on average (38.83% for solidity, 32.90% for support). The only difference between the methods used in the pilot study and in Study 1 is that the pilot was conducted entirely between subjects, and Study 1 used a within-subjects design, where all infants saw both expected *and* unexpected events. Infants were divided evenly into the two main conditions (solidity and support). With 30 infants in each condition, we would have 99.7% power to detect a difference in looking time between expected and unexpected events in the solidity conditions (where Cohen's d = 1.24) and 100% in the support conditions (where Cohen's d = 2.06). We recognize, however, that with such a small sample size in our pilot sample, the effect sizes were likely exaggerated. Thus, we also verified that 30 infants per condition would allow us to detect an effect size that is large but not as large and found that we would have 76% power to detect a large effect (Cohen's d = .7). Though we shifted from a between-subjects to within-subject design, we did not expect effect sizes to get smaller. Therefore, we proceeded with 30 infants per condition.

Design

All participants were tested live online over Zoom. The experimenter and the participant were seated in front of their screens so that they had clear visual access to the shared display. Each infant saw either solidity events or support events. They first saw two identical familiarization trials followed by four event trials (two that should have been expected based on principles of solidity or support, and two that should have been unexpected). Thus, there were two main conditions: solidity and support. Events were presented in A-B-B-A order, such that if an infant first saw an expected event, they would then see two unexpected events, and finally another expected event. Order of events trials was counterbalanced, such that half of infants saw an unexpected event first and the other half saw an expected event first.

Infants were randomly assigned to one of the conditions. In order for the experimenter to remain blind to the condition, conditions were randomly renamed, and the monitor displaying the stimuli to the infant were never visible to the experimenter.

Procedure

Parents were instructed to seat their child (either in a highchair or in their lap) as close to the computer monitor as possible without the infant being able to reach the computer or keyboard. Parents were also instructed not to interact with their infant or to direct their infant's attention throughout the duration of the study. After informing and instructing parents, infants were shown a series of calibration animations to provide a reference point for what it looked like as they attended to different locations on the screen, as each child's monitor setup and orientation was unique.

The experimenter covered her screen in such a way that she had access to the live video of the infant watching the display but not to the specific stimuli the infant saw. The videos had sound cues that informed the experimenter of when the trial was starting, when the outcome had been revealed, and when the trial had ended. Sound cues were identical between expected and unexpected events, so that the experimenter remained blind to trial order. If the experimenter saw that the infant was not looking during crucial portions of the trial (based on these sound cues), she started the trial over to ensure infants saw the full events and outcome.

<u>Solidity condition</u>: In familiarization trials, infants watched as a hand reached in from above and placed an occluder in front of the right side of the stage. A hand then reached in from

above holding a ball and waved it for 3 seconds while a voice said, “Hi baby! Look at this!”. The ball was then placed at the top of the ramp and was released. It rolled down the ramp, across the stage, and behind the occluder. A hand then reached in from above and removed the screen to reveal that the ball had come to rest at the far right side of the stage (Figure 2.1, panel a), and this final position remained on screen for 5 seconds. This exact trial was repeated a second time. The subsequent event trials were identical to the familiarization trials with two exceptions; (1) before the occluder was placed, the experimenter reached in from above holding a pink wall. She

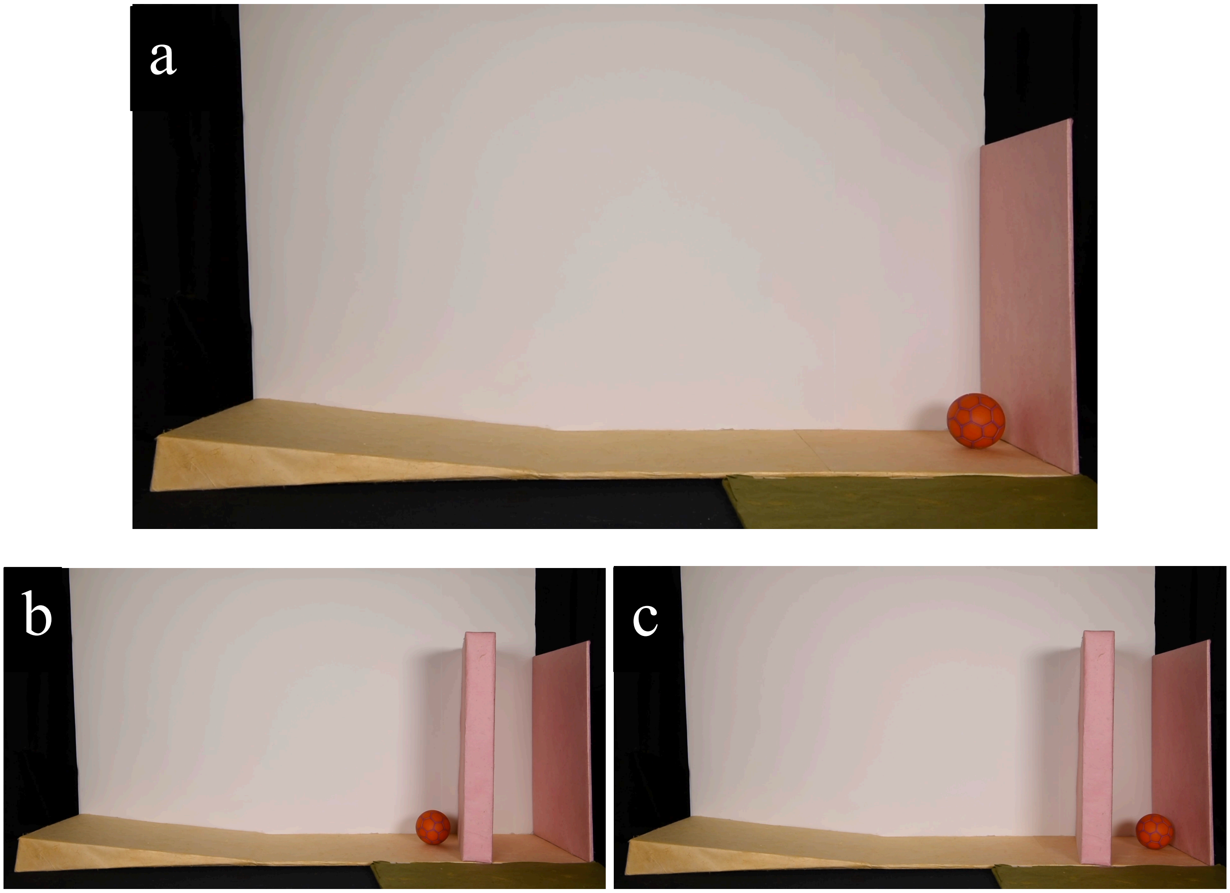

Figure 2.1 Stimuli for solidity events in Study 1, including (a) the familiarization event, where the ball is revealed at the base of the rightmost wall, (b) the expected test event, in which the ball is revealed to the left of the blocking wall, appearing to have stopped when it hit the wall (thereby adhering to the principle of solidity), and (c), the unexpected test event, in which the ball is revealed between the blocking wall and the rightmost wall, appearing to have rolled through the blocking wall (thereby violating the principle of solidity).

knocked on it twice and then turned it and placed it across the center of the track near the rightmost wall; (2) When the occluder was removed at the end of the trial, the ball remained in its final position for 30 seconds. For expected trials, the final position of the ball was to the left of the pink wall (Figure 2.1, panel b), and for unexpected trials it was to the right of the pink wall (Figure 2.1, panel c).

Support condition: In familiarization trials, infants watched a hand reach in from the left side of the screen holding a ball. The hand waved the ball just above the box for 3 seconds while a voice said, “Hi baby! Look at this!”. The hand then placed the ball atop a striped box midway between the left and right edges of the box. The hand again reached in from the left side, pushed the ball to the rightmost edge of the box (Figure 2.2., panel a), then returned to its original

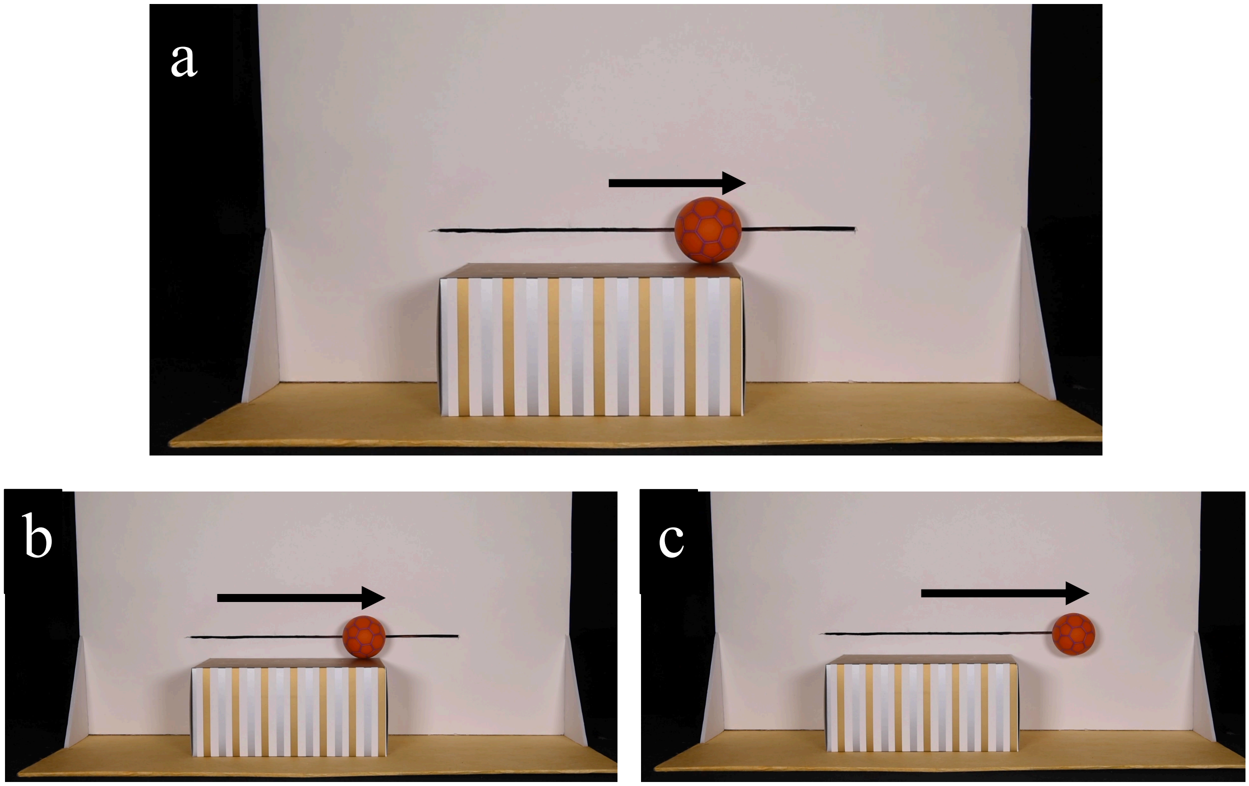

Figure 2.2 Stimuli for support events in Study 1, including (a) the familiarization event, where the ball is placed at the midpoint of the box and is then pushed to the rightmost edge, (b) the

expected test event, in which the ball is placed at the leftmost edge of the box and is pushed to the rightmost edge of the box, remaining in contact with the box throughout (thereby adhering to the principle of support), and (c), the unexpected test event, in which the ball is placed at the midpoint of the box and is pushed passed the edge of the box, appearing to hover (thereby violating the principle of support). Arrows represent movement of the object in each trial.

position (out of view). The ball remained visible in its final position for 5 seconds before the trial ended. This exact trial was repeated a second time. The subsequent event trials were identical to the familiarization trials with one exception: the ball was pushed twice as far. For expected trials, the ball was placed at the far left side of the box and was pushed to the far right side of the box (Figure 2.2, panel b), and unexpected trials, it was placed in the center of the box and pushed several inches past the end of the box (Figure 2.2, panel c).

Coding and analyses

All looking time was coded offline once data collection was complete. A coder who was unaware of the hypotheses reviewed the videos to ensure that infants met inclusion criteria, which had been set prior to data collection. Specifically, trials were excluded if (1) there were technological errors, including infants pressing buttons that caused the display to exit full-screen mode (n = 2), the video froze for more than two seconds (n = 1), or Zoom displayed a pop-up notification during the recording of looking time (n = 1), (2) infants were not looking during a crucial part of the trial, specifically when the object rolled across the stage in the solidity event or when it was pushed across the box in the support event (n = 1), or did not look within the first two seconds of the reveal of the outcome (n = 2), or (3) the infant moved such that their eyes were out of view for more than two seconds and looking could not be clearly inferred from head position (i.e. if their head was turned away) (n = 1). When a trial was excluded, the complimentary trial in the set was also excluded, so that the infant contributed only one expected and one unexpected trial. In other words, if either of the first two trials met exclusion criteria the other of the first two trials was also excluded, and data from the third and fourth trial were included (and vice versa if either the third or fourth trial was excluded).

In all test trials, the coder began recording looking time as soon as the object was fully revealed in its final position and stopped either when the infant looked away for 2 consecutive seconds or at the end of the 30 second trial. All trials were coded without knowledge or visual access to trial type or condition. To assess reliability, a second coder who was also blind to condition coded a random 25% of videos. The intraclass correlation (ICC) between the two coders was 0.94, 95% confidence interval [0.894, 0.965]. Given this high degree of agreement, the primary coder's codes were used for analyses.

Inferential statistics were fit to log-transformed looking times, as the log-normal distribution provided a better fit than did the normal distribution (log-likelihood = -717.10 as compared to -736.59)[2]. We used linear mixed effects models to conduct ANOVAs and paired t-tests in R (version 4.2.1; RCore Team, 2013), using the lme4 package (Bates et al., 2014). Because participants were tested across multiple trials, all models included subject identity as a random effect in order to account for repeated measures. We also used Cook's Distance to determine whether there were any participants whose data disproportionately impacted the results using the influence.ME package (Nieuwenhuis et al., 2012). We used ggplot2 to plot all data (Wickham, 2016). We fit three categories of models for this data, (1) a null model with subject identity as the only factor, (2) hypothesis-driven models preregistered on Open Science Framework, and (3) exploratory models, which included additional factors that were not hypothesis-driven. We used Akaike information criteria (AICs) to evaluate model fit.

---

[2] Note that plots and descriptive statistics still use raw looking times so that they are easier to interpret.

2.2.2 Results

Hypothesis-driven results

Our hypothesis driven model included event type (solidity or support) as a between-subjects factor and outcome type (expected or unexpected) as a within-subjects factor, with subject identity as a random effect. This analysis revealed that looking time did not differ depending on whether infants were viewing solidity events (M = 9.46 s per trial, SD = 7.20) or support events (M = 8.96 s per trial, SD = 5.39) 95% confidence interval (CI) [-0.304, 0.336], b coefficient (B) = 0.016, standard error (SE) = .160, t(50) = 0.102, p = 0.919, two-tailed, nor did it differ depending on whether the event violated a principle of core knowledge (M = 9.07 s per trial, SD = 6.16) or it did not (M = 9.33 s per trial, SD = 6.49), 95% CI [-0.161, 0.184], B = 0.011, SE = .087, t(154) = 0.126, p = 0.900, two-tailed.

An analysis using Cook's Distance (4/*n*, where *n* is equal to the number of subjects; Van der Meer et al., 2010) indicated that three participants were especially influential (cutoff = 0.066, D = 0.087, 0.068, 0.200). Excluding these participants did not significantly impact the output of the hypothesis-driven model, overall 95% CI [1.794, 2.199], B = 1.996, *SE* = 0.102; fixed effect of event type 95% CI [-0.231, 0.291], B = .003, SE = .131, t(55) = 0.23, p = 0.819, two-tailed; fixed effect of trial type 95% CI [-0.158, 0.170], B = 0.006, SE = 0.083, t(160) = 0.07, p = 0.945, two-tailed.

Subsequent planned paired t-tests revealed no significant difference in looking time between expected versus unexpected trials for either solidity events, 95% CI = -0.246, 0.238], B = -0.004, SE = 0.122, t(73) = -0.033, p = 0.487, one-tailed. nor for support events, 95% CI [-0.223, 0.272], B = 0.025, SE = 0.125, t(80) = 0.199, p = 0.422, one-tailed.

An analysis using Cook's Distance for the model evaluating looking to solidity events indicated that one participant was especially influential (cutoff = 0.133, D = 0.149). Excluding this participant did not significantly impact the output of the hypothesis-driven model, 95% CI [-0.246, 0.246], B = -0.0003, SE = 0.124, t(73) = -0.003, p = 0.499, one-tailed. Such an analysis for the model evaluating looking to support events indicated that four participants were particularly impactful (cutoff = 0.129, D = 0.145, 0.147, 0.163, 0.210). Excluding these

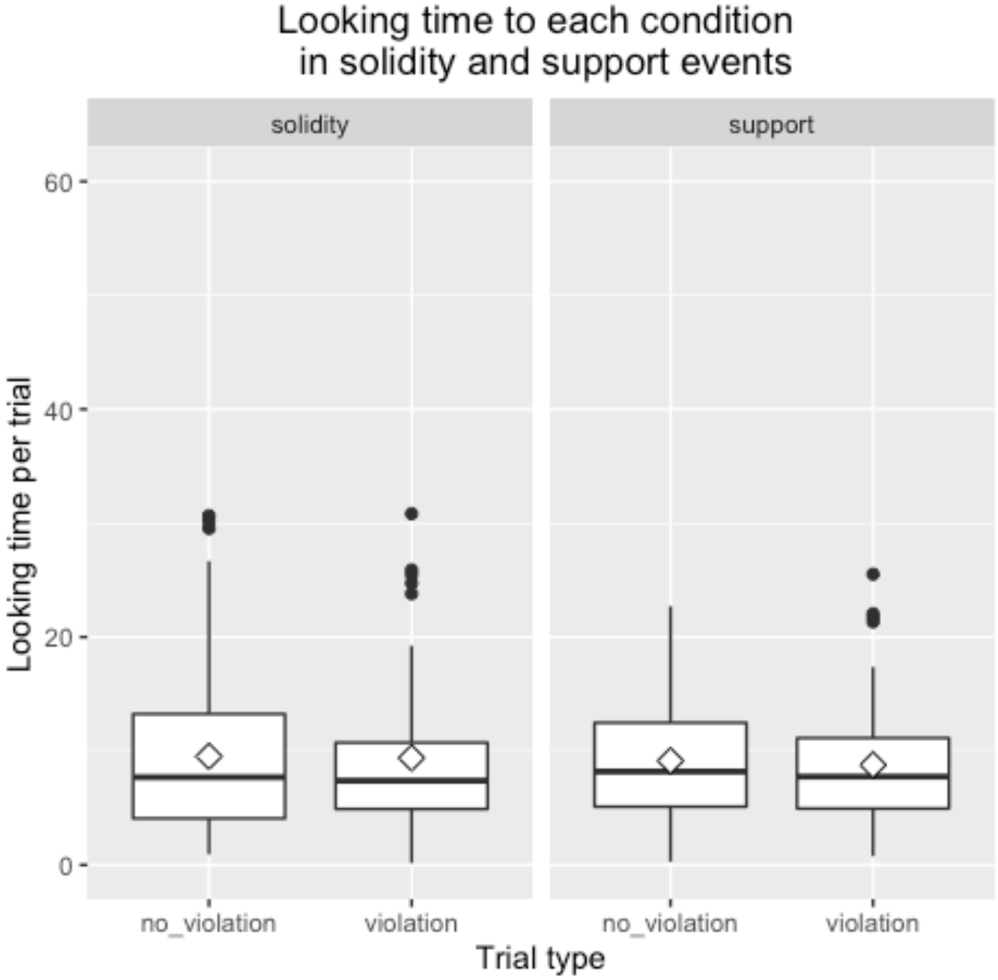

Figure 2.3 Boxplots for raw looking time (in seconds) to test events in Study 1 (n = 32 per event type). Event type (solidity vs support) was between subjects while trial type (solidity/support violation vs no violation) was within subjects. Boxes demarcate the interquartile ranges, bolded lines demarcate the medians, and white diamonds demarcate the means.

participants did not significantly impact the output of the hypothesis-driven model, 95% CI [-0.200, 0.183], B = -0.010, SE = 0.097, t(79) = -0.093, p = 0.463, one-tailed.

Exploratory results

An exploratory model that additionally assessed age and sex as fixed effects revealed that looking time did not significantly either age, 95% CI [-0.048, 0.570], B = 0.260, SE = 0.155, t(48) = 1.680, p = 0.099, or sex, 95% CI [-0.275, 0.354], B = 0.042, SE = 0.157, t(49) = 0.266, p = 0.791. Another exploratory model that included trial type (expected vs unexpected), trial order (whether an expected or an unexpected event was displayed first), and their interaction yielded no significant order effects 95% CI [-0.548, 0.584], B = -0.096, SE = 0.175, t(153) = -0.548, p = 0.584.

An additional model including only trial order as a fixed effect and subject identity as a random effect revealed a trend of looking time decreasing across trials (Figure 2.4), with a statistically significant drop-off by Trial 4, 95% CI [-0.527, -0.042], B = -0.284, SE = 0.123, t(159) = -2.310, p = 0.022, two-tailed. We also conducted the hypothesis-driven model on just the first two trials, as attention seems to have been a bit higher during these trials, and still there was no effect of event type, 95% CI [-0.590, 0.120], B = -0.235, SE = 0.178, t(59) = -1.318, p = 0.193, two-tailed, or trial type. 95% CI [-0.127, 0.375], B = 0.124, SE = 0.126, t(59) = 0.985, p = 0.329, two-tailed. We then conducted this same model again on just the last two trials, as differential looking may have been revealed once the events were less novel, but again no effect was observed either for event type, 95% CI [-0.045, 0.659], B = 0.307, SE = 0.176, t(54) = 1.743, p = .087, two-tailed, or for trial type, 95% CI [-0.321, 0.096], B = -0.113, SE = 0.105, t(54) = -1.077, p = 0.286, two-tailed.

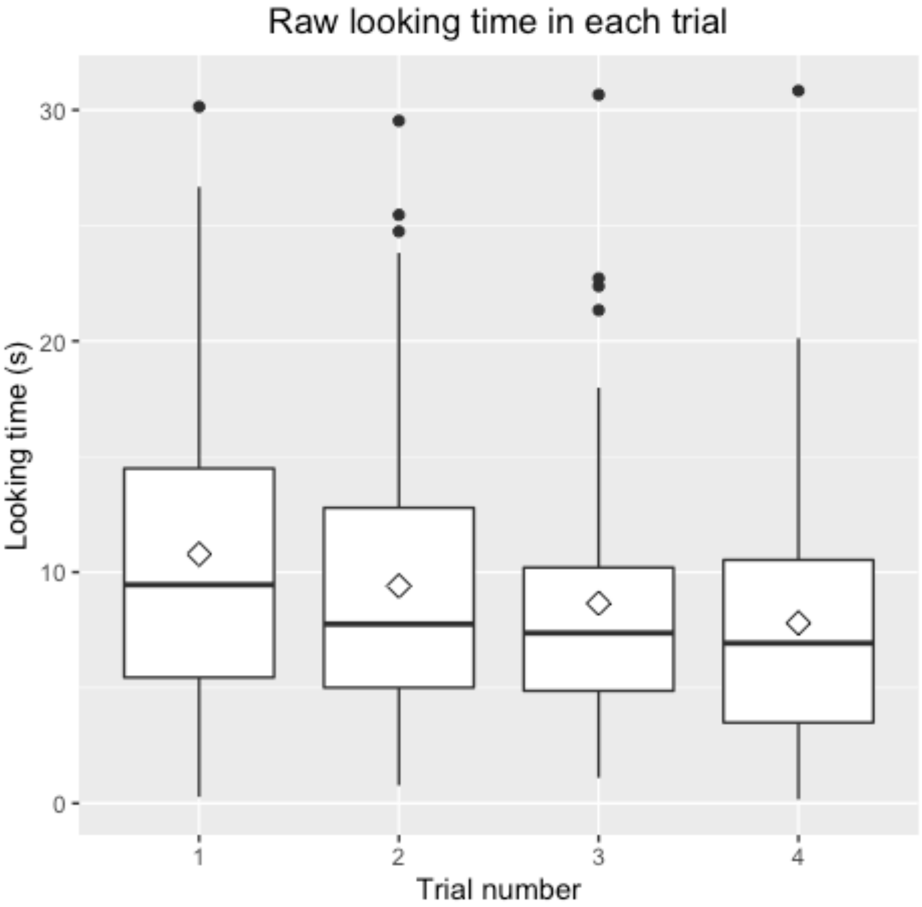

Figure 2.4 Boxplots for raw looking time (in seconds) in each trial, combined across event types to demonstrate trial effects on looking time (n = 64). Boxes demarcate the interquartile ranges, bolded lines demarcate the medians, and white diamonds demarcate the means.

2.2.3 Discussion

This study revealed that, contrary to previous results using VOE methods, infants did not differentially attend to events depending on whether or not they had violated the core knowledge principles of solidity or support. As revealed by our hypothesis-driven models, infants who saw solidity events looked for approximately the same duration overall as did infants who saw support events. Regardless of what types of events they saw, infants did not look significantly longer when viewing a violation event. Thus, infants in this study did not show a VOE effect when witnessing a violation of solidity or of support.

Our exploratory analyses showed that looking time did not differ with age or sex, nor did it differ depending on whether infants first saw an event that violated or adhered to principles of solidity or support. There was a decline in looking overall across events, and by the fourth trial, looking had significantly decreased from the start of the test trials. Further analyses of looking in just the first set of trials and then in just the second set of trials still did not reveal any difference in looking depending on whether or not the event violated a principle of core knowledge. As the null model provided the best fit, it seems that none of the factors we examined accounted for variation in looking time across infants or trials.

There are a number of potential explanations for why we did not see infants tending to look longer to the surprising events. One possibility is that the specific videos that we created to depict violations of solidity and support were in some way flawed such that infants were unconvinced by the violation. We find this to be an unlikely explanation, as the videos were modeled directly after events used by Stahl & Feigenson (2015), and many previous studies used very similar displays (eg: Baillargeon et al., 1992; Spelke et al., 1992).

A related possibility is that infants were able to use cues in the scene to provide themselves with explanations for the surprising events, thus making them unsurprising. Baillargeon (1994) posits this explanation for a series of studies in which infants did not look longer to events that were designed to violate their expectations. For example, infants were shown two dolls being hidden behind a wall, and three dolls were then revealed. Surprisingly, infants did not look for a significantly longer time than if two dolls were then revealed. Baillargeon proposed that infants might not find this discrepancy surprising, because they may have generated the explanation that there was initially a doll behind the wall before the two were hidden there. To test this, she ran another study where some infants had visual access to the space behind the wall before it was positioned. Infants who were shown that there were no dolls behind the wall now looked longer when two dolls were hidden and three were revealed, as compared to infants who did not get to see behind the wall. It is therefore plausible that infants in Study 1 were similarly positing an explanation that would render the surprising events to actually be unsurprising. For example, in the support events, we achieved the visual effect of the object floating in midair by attaching a stick to the back of the object which passed through the back of the display. The track for this stick was visible, so it is theoretically plausible that infants perceived the object to be moving along this line and therefore provided themselves with an explanation for the object hovering.

A third possibility is that we failed to find a VOE effect because stimuli were representations of a 3-dimensional scene presented on a 2-dimensional display. Past work has revealed a number of obstacles to conducting research via a 2-dimensional display, other researchers who study surprise have similarly struggled to elicit surprise-related learning effects using screens to display the stimuli (eg: Stahl, personal communication). Research has demonstrated that certain visual cues must be present in order for infants to show VOE effects to video as they would to real-life displays, such as sufficient background texture (Johnson & Aslin, 1996). There is further evidence that 9-10-month-old infants prefer to attend to, attend longer to, and react more to live as compared to video displays of the same evidence (Diener et al., 2008). Infants also have some difficulty transferring information learned from a 2-dimensional format (either a picture book or a screen) to the real world, including learning a demonstrated action or a new word (see Barr, 2013 for review; Barr & Hayne, 1999). Additional studies have revealed that even preschoolers still struggle to map information taught on a 2-dimensional display onto the real-world scene that it represents. 2-2 ½ year-olds watched as an adult hid an object in a room. Some participants watched the demonstrator in person and others were shown over a digital display with extensive orienting information about the relationship between the room on the screen and the testing room. 2-year-olds failed to use the information when it was provided on a screen, but by 2 1/2, children seemed better able to utilize this information (Troseth & DeLoache, 1998). A replication of this study found that even by 2 ½ children are still performing significantly better when they see the experimenter hide the object in real life, and it is not until children are 3 years old that they can reliably locate the object based on the information they were given virtually as well as if it were presented live (Schmitt, 1997). Troseth & DeLoache (1998) suggest that it is not visual cues that make these tasks difficult, but rather that it is difficult for infants and children to perceive images on a screen as a representation of the real world. They provided evidence for this proposal by conducting the previously described study with an additional condition in which they led children to believe that the screen they were watching was a window. Children performed better in this condition, providing some evidence that it is not merely the visual properties of a screen that virtual tasks more challenging.

The problem of evoking VOE to virtual stimuli may be particularly profound in today's world, where infants have abundant exposure to 2-dimensional displays that make them unreliable as accurate representations of the 3-dimensional world. Principles of core knowledge constantly appear to be violated on these screens; object continuity appears to be violated every time we swipe from one image to the next on a smart phone, objects often appear to hover unsupported on a screen, etc. Thus, the infants may not have looked longer to the ball rolling through the wall or hovering in midair in Study 1 because they do not hold the same expectations of objects when they are on a screen as when they are interacting in the real 3-dimensional world. Due to limitations imposed by the COVID-19 pandemic, we were unable to directly test this possibility by conducting parallel studies with video and live-display stimuli, as planned.

Contrary to this proposal though, Smith-Flores et al. (2022) found evidence of 15-month-olds looking longer to violations of support in a recent online study.

Another possibility is that infants at this age do not show a VOE effect, despite holding expectations of object solidity and support. Previous work was largely conducted with younger infants who are young enough that their primary form of exploration is visual. It is possible that older infants, who can more actively engage with their surroundings, do not engage in longer looking. The Smith-Flores et al. (2022) study described above contradicts this explanation, though it could still be the case that looking effects are less robust at this age.

Without further in-person manipulations, it is difficult at this point to identify the precise cause of our failure to detect a VOE effect to solidity and support violations in 11-month-old infants. What is apparent here is that looking time to these events may not be as sensitive and robust a measure as previously proposed. However, based on the myriad studies that preceded Study 1, we find it exceedingly unlikely that infants do not reason based on core knowledge principles of solidity and support. Instead, we merely demonstrate that VOE measures might not always accurately reflect these expectations across settings, display media, and age.

2.3 Study 2: Macaques

2.3.1 Methods

Subjects

Study 2 included 80 monkeys, 40 juveniles (ages 2-5 years) and 40 adults (older than 5 years) (20 females in each age group, mean = 7.0 years; age range = 1.7-21.5 years). An additional 50 monkeys were tested but were not included in the final sample due to experimenter or stimulus error as a result of the uneven terrain (n = 13), the participant leaving partway through testing (n = 20), interference by another monkey (n = 5), the participant losing interest and failing to orient to the apparatus (n = 5), the monkey looking away for a crucial part of the trial (n = 3), and errors with the video that prevented accurate coding, such as the monkey's eyes leaving the video frame (n = 4).

All data was collected at Cayo Santiago, was approved by the Cayo Santiago IACUC, and adhered to all guidelines as such. Cayo Santiago is home to approximately 1,000 semi-free ranging rhesus macaques. The monkeys have had exposure to researchers and other humans all their lives and are thus relatively habituated to human presence. All monkeys are marked with individual identifiers, which allowed us to ensure the same participant was not tested twice and also allowed for accurate age identification. As monkeys freely range on the island, the experimenter tested monkeys in their home environment. Monkeys would be eligible to be tested if they were sitting relatively isolated from other adults and were positioned in such a way that the experimenter could place the apparatus between themselves and the monkey at an appropriate distance, and at an angle at which the object could properly roll across the apparatus.

Design

When testing a participant, the experimenter knelt approximately 2 meters in front of the monkey with the apparatus placed directly in front of the experimenter, facing the monkey. A second experimenter video recorded the participant's face from behind the primary experimenter.

The apparatus was analogous to that used in the solidity events in Experiment 1. In this case, it was a white stage with a white panel on the left side, a black panel on the right side, and a second, identical black panel several inches from the right edge of the stage. A white ramp sloped from the left panel, such that an object placed on the left side of the stage would roll towards the right side of the stage. A colored panel served as an occluder, such that it could be raised to cover the right half of the stage or lowered to reveal it. When the occluder was raised, the vertical panels both remained visible over the top of the occluder. The back of the apparatus had two holes, one at each of the two test event outcome locations, such that objects could be placed or removed through the back of the apparatus while the occluding wall was in place.

All monkeys were first shown one familiarization trial to habituate them to the scene and

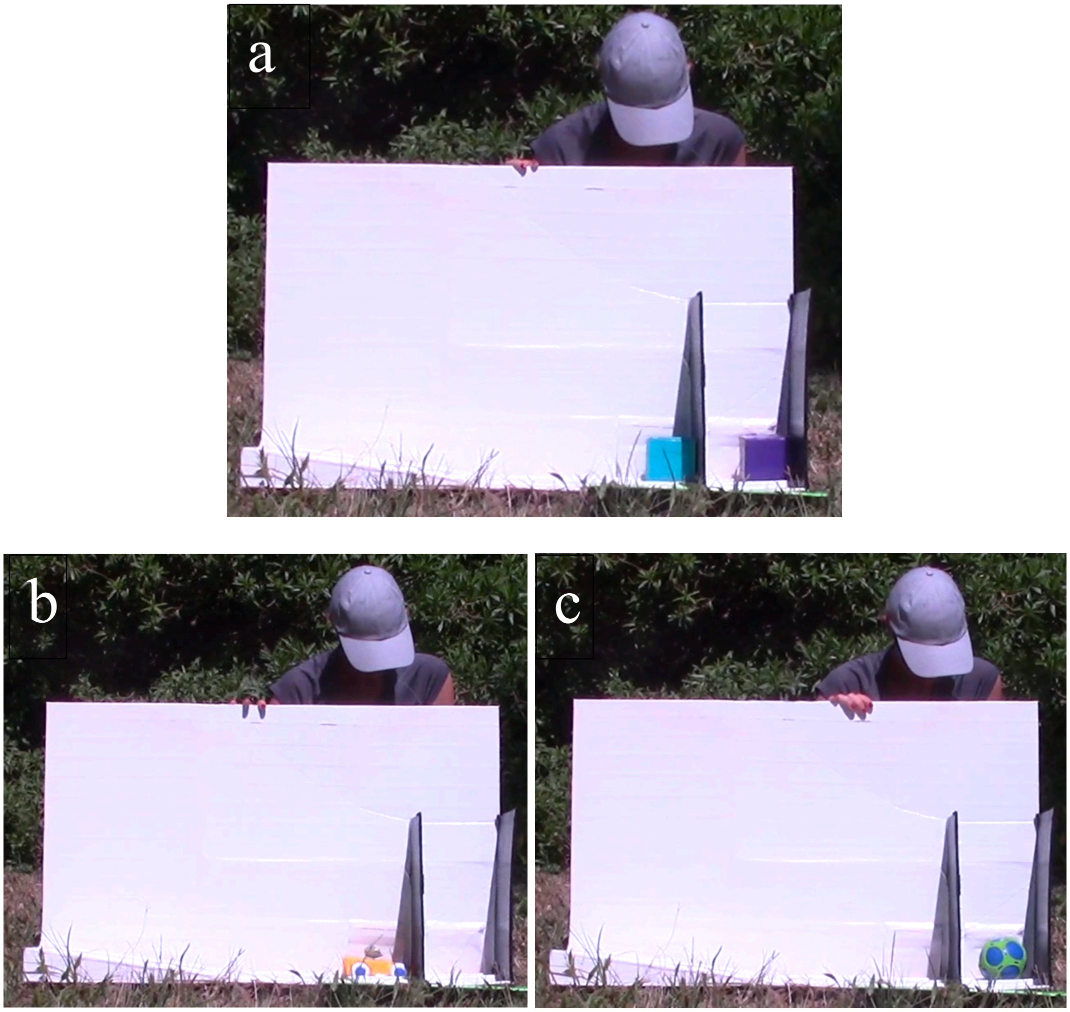

Figure 2.5 Stimuli for Study 2, including (a) the familiarization event, where the objects are revealed at the two possible locations, (b) the expected test event, in which the object is revealed to the left of the blocking wall, appearing to have stopped when it hit the wall (thereby adhering to the principle of solidity), and (c), the unexpected test event, in which the object is revealed between the blocking wall and the rightmost wall, appearing to have rolled through the blocking wall (thereby violating the principle of solidity).

to seeing objects at each of the two potential outcome locations so as to dampen the effect of participants looking more to the first trial (Figure 2.5, panel a). They were then shown two event trials, one expected (Figure 2.5, panel b) and one unexpected (Figure 2.5, panel c). In the expected event, the monkeys saw an object roll down the ramp, and when the occluder was lowered, the object was in front of the first barrier (L1), as would be expected based on the principle of solidity. The unexpected event is identical except that when the occluder was lowered, the object was in front of the second barrier (L2). Order of trials and which object was used in which event was counterbalanced.

Procedure

Once the experimenter was situated in front of the participant, she tapped on the occluder and called "monkey" to draw the participant's attention to the right side of the apparatus. She then lowered the occluder to reveal the right side of the stage and knocked twice on the black barrier wall to draw monkeys' attention to the wall and to demonstrate its solidity.

For the familiarization trial, the experimenter placed two colored cubes, one at L1 and one at L2, while the occluder was raised. She then called the monkey's attention to the right side of the stage, lowered the occluder, and left the scene fully visible for 10 seconds while looking time was recorded. She then reached down and removed the two objects and raised the occluder. For the event trials, the experimenter held the object (either a yellow car or a blue and green ball) at the top of the ramp and drew the monkey's attention. She then released the object so that it rolled down the ramp, across the stage, and behind the occluder. She then removed the object from the flap at L1 and either replaced it at L1 (expected trial) or replaced it through the flap at L2 (unexpected trial). Until this point, the main experimenter was unaware of which condition she would run, so as to remain as unbiased in the stimulus presentation as possible; the secondary experimenter told the primary experimenter at this point whether to replace or move the object before revealing the outcome to the participant. She then removed the object, raised the occluder, and repeated the trial with the other object and the opposite outcome. All outcomes were visible for 10 seconds. The experimenter wore a hat with a brim and looked down throughout all periods in which looking time was recorded, in order to prevent her from influencing monkeys' gaze direction.

The primary experimenter made the real-time decisions of when to abort a session because she was blind to condition and to the monkeys' responses. Aborted trials were typically called when a monkey approached or made contact with the apparatus, another monkey interfered with the trial, or the subject walked away.

Coding and analyses

All looking time was coded offline once data collection was complete. Videos were reviewed regularly throughout data collection to ensure that monkeys met inclusion criteria, which had been set prior to data collection. When a monkey was excluded, an additional monkey was tested in the same condition in order to reach the target sample size after exclusions.

In all test trials, the coder began recording looking time as soon as the experimenter lowered the occluder such that the object was fully revealed in its final position, and looking time was recorded for 10 seconds. Videos were filmed to include only the monkey's face and exclude all information about the condition, and all clips were randomized, so coders were able to evaluate looking time without any knowledge of the trial outcome. To assess reliability, a second coder who was also blind to condition coded 100% of videos. The intraclass correlation

(ICC) between the two coders was 0.975, 95% confidence interval [0.966, 0.982]. Given this high degree of agreement, the primary coder's codes were used for analyses.

Inferential statistics were fit to raw looking time, as the log-normal distribution provided a worse fit than did the normal distribution (log-likelihood = -378.96 as compared to -372.53). We used linear mixed effects models in R (version 4.2.1; RCore Team, 2013), using the lme4 package (Bates et al., 2014). We also used Cook's Distance to determine whether there were any participants whose data disproportionately impacted the results using the influence.ME package (Nieuwenhuis et al., 2012). As in Study 1, we then conducted a null model, a hypothesis-driven model, and exploratory models. The remainder of analysis methods were the same as those used in Study 1.

2.3.2 Results

Hypothesis-driven results

The hypothesis-driven model, which included condition (solidity violation vs no violation), age, and their interaction as fixed effects, and subject identity as a random effect, revealed a main effect of age, such that as age increased looking time decreased (Figure 2.6), 95% CI [-0.266, -0.062], B = -0.164, SE = 0.052, t(153) = -3.134, p = 0.002. There was no significant effect of condition (Figure 2.7), 95% CI [-.538, 1.787], B = 0.624, SE = 0.594, t(77) = 1.052, p = 0.296, or the interaction of age and condition, 95% CI [-0.218, 0.055], B = -0.081, SE = 0.069, t(77) = -1.173, p = 0.244. An analysis using Cook's Distance indicated that one participant was especially influential (cutoff = 0.05, D = 0.08). Excluding this participant did not significantly impact the output of the hypothesis-driven model. A likelihood ratio test (LRT) demonstrated that the hypothesis-driven model provided a better fit than did the null model that included only subject identity as a factor $x^2(3) = 16.68$, $p = 0.0008$.

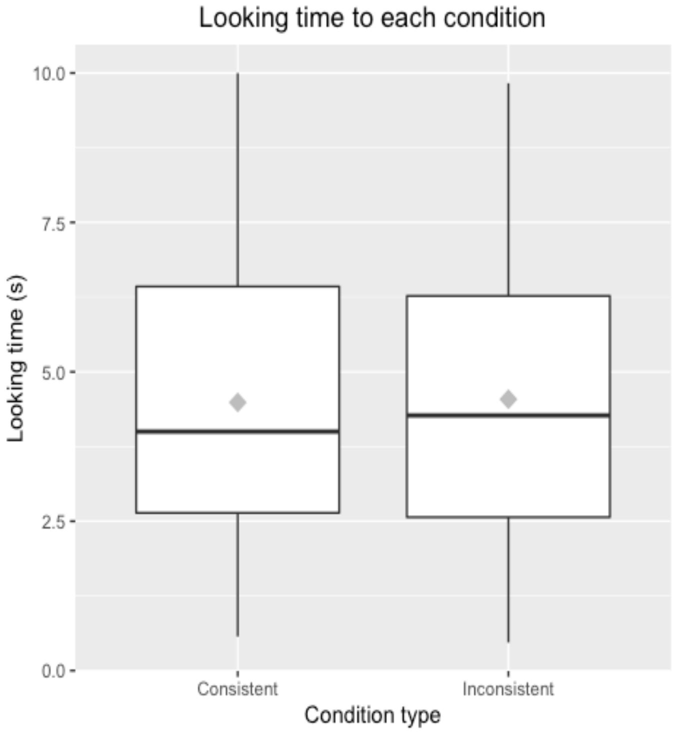

Figure 2.7 Boxplots for raw looking time (in seconds) to test events in Study 2 (n = 80), which were consistent or inconsistent with the principle of solidity. Boxes demarcate the interquartile ranges, bolded lines demarcate the medians, and grey diamonds demarcate the means.

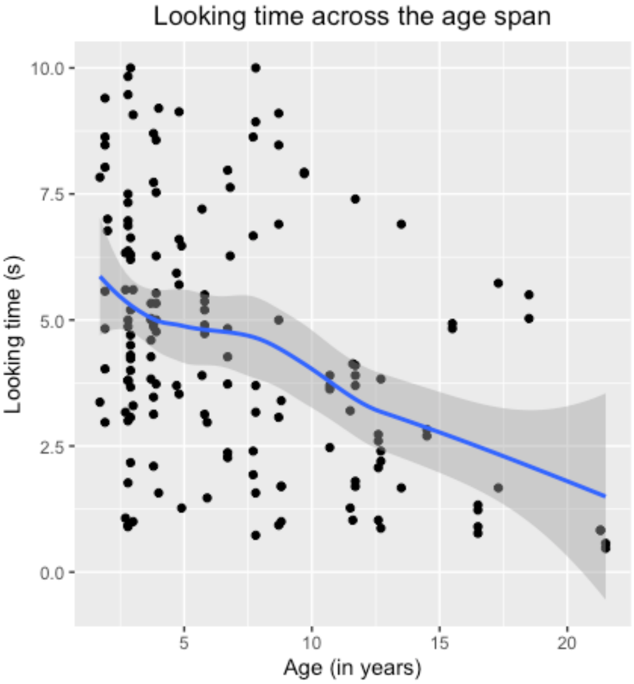

Figure 2.6 Raw looking time per trial across the age span. Points are individual trials, the blue line indicates the trend, and the gray shadow indicates the standard error

Exploratory results

An exploratory model, which additionally included sex, trial order, object order, and which object participated in the violation event revealed that only sex predicted looking time, such that males had higher looking times overall than did females 95% CI [-0.259, -0.058], B = 0.798, SE = 0.383, $t(148) = -3.015$, $p = 0.041$. An LRT showed that this exploratory model did not provide a better fit than did the hypothesis-driven model $x^2(4) = 4.621$, $p = 0.329$.

2.3.3 Discussion

These results demonstrated rhesus macaques in our sample did not look longer when an object appeared to roll through a wall than when it appeared to stop at that wall. While this result failed to support past findings of VOE to solidity violations in primates (eg: Santos & Hauser, 2002; Cacchione et al. 2009), it does not provide conclusive evidence to contradict that monkeys reason according to the core knowledge principle of solidity and find it surprising when an object appears to violate this principle.

While looking time decreased overall across the age span, there was no evidence that differential looking between solidity-consistent and solidity-inconsistent events varied across the age span. Thus, younger monkeys seemed as to be as likely as older monkeys to look longer at the unexpected event, indicating that our failure to detect longer looking to the solidity violation is not due to the expanded age range of the present study. Because we did not successfully elicit longer looking, and because we were unable to test infant monkeys that have not yet been individually identified, subsequent studies should continue to investigate the developmental trajectory of solidity expectations in primates across the lifespan in order to better understand the role of experience.

We propose several explanations as to why there was not a statistically significant difference in looking time across conditions. The first possibility is that monkeys were not as deeply encoding the events as in past studies because the objects used in our study were not as ecologically valid as were the food items used in past work (eg: Santos & Hauser, 2002). While monkeys may have the ability to reason about inanimate and inedible objects, the lack of relevance to their lives (at present and evolutionarily) may lead them to be less readily represented. Particularly in an uncontrolled environment, where distractions are plentiful, an inanimate object may not have held the monkeys' full concentration and attention for deeper encoding.

A second possibility is that the monkeys were sensitive to cues about the manipulations being performed on the apparatus and that this therefore made the surprising event less convincing. For one, during the familiarization trial, the objects were introduced through the trap doors, rather than being placed in view of the monkey. For another, monkeys may have detected auditory or visual cues as the object was being removed and either moved or replaced that may have dampened any surprise effect. A follow-up study where the familiarization objects were placed in view of the monkey and the apparatus was designed to better facilitate movement of the objects could help to address these possibilities.

2.4 General discussion

While the findings discussed in this chapter do not necessarily support previous findings, they are still consistent with the proposal that both human infants and primates share foundational knowledge about how objects behave. These studies indicate that VOE studies may not be as sensitive and robust as they are often discussed to be, but it does not undermine decades of research demonstrating that infants and primates are sensitive to such violations. Study 1 points to the need for further research investigating infants' representations of scenes depicted on 2-dimensional displays and infants' abilities to explain away the surprising events they are shown, but it does not demonstrate a resounding lack of core knowledge of object solidity and support in 8-10-month-olds. The findings from Study 2 suggest that further work might elucidate primates' sensitivity to violations of core knowledge of physics depending on the relevance of the event to the individual, such as whether it involves food, or whether they might otherwise benefit from devoting attention to the event.

Ultimately, we propose that these findings call into question the sensitivity of looking time as a measure to gain access into individuals' understanding of object relations, but not that they challenge the existence of such understanding.

Chapter 3 Bonobos' use of core knowledge to reason about the location of hidden food

3.1 Background

Individuals are constantly faced with the task of reasoning about inanimate entities, including their properties, movements, and interactions with other objects and their environments. Imagine an ape is sitting on a wide tree branch foraging for fruit, and as they go to place a morsel into their mouth, they accidentally drop it. Where will the ape search for this released food? Will they look in the exact spot where it last had contact with their hand, expecting it to remain stationary in midair? Will they look beneath the branch on which they sit, expecting that it fell through their arboreal perch? Or will they correctly deploy principles of core knowledge of physics, recognizing that the food would fall if unsupported but would stop when it came into contact with the solid branch?

Prior work has indicated that primates do indeed have expectations of objects that are consistent with core knowledge. Evidence of such expectations have come from studies that use the more passive measure of how long individuals' look at different events as an indication of what they expect, as well as from a handful of studies that use more active measures, such as pointing and searching.

3.1.1 Primates' core knowledge of objects: Evidence from looking

As discussed in detail in Chapter 2, studies with both monkeys (Hauser & Carey, 2003; Munakata et al., 2001; Murai et al., 2011; Santos & Hauser, 2002; Uller et al., 1997, 2001) and apes (Cacchione & Krist, 2004; Murai et al., 2011) have demonstrated sensitivity to violations of core knowledge of objects using looking time. Santos & Hauser (2002) found that rhesus macaques (*Macaca mulatta*) look longer when an object appears to roll through a wall or fall through a shelf than when it appears to stop when it hits the wall or the shelf, demonstrating evidence of expectations about object solidity. Murai et al. (2011) provided evidence that Japanese monkeys (*Macaca fuscata*) reason about support, as they looked longer when an object was pushed passed the edge of a surface or had only minimal contact with a surface but remained floating (though interestingly, they did not look longer when the object had contact only with the side of the surface, which should actually not provide sufficient support). This same pattern of looking has also been demonstrated in chimpanzees (*Pan troglodytes*) (Cacchione & Krist, 2004; Murai et al., 2011), so they too seem to recognize violations of object support.

As demonstrated and discussed in Chapter 2, the violation of expectations (VOE) method is not necessarily always a sensitive, robust, or ecologically valid measure for assessing underlying expectations in primates. Looking time is certainly a valuable tool, as it is nonverbal, requires little to no training, and can be used flexibly across settings. However, it requires individuals to maintain attention to a display with limited interaction and typically no direct rewards. While absence of evidence certainly should be interpreted with an abundance of caution, it is worth noting that the study of VOE to support events in chimpanzees (first conducted by Cacchione & Krist, 2004 and subsequently replicated by Murai et al., 2011) is the only published study of which we are aware that utilized VOE to demonstrate expectations about objects in apes. In our own attempts to execute this method with bonobos (*Pan paniscus*), we found that individuals were highly inattentive to the display and often did not wish to participate, even when food rewards were provided between trials to incentivize attention.

Eye-tracking methods have yielded greater success in measuring apes' attention to different stimuli (see Hopper et al., 2021; Lewis & Krupenye, 2022 for reviews). In these studies, apes are typically positioned with a juice mount in front of a screen that is equipped with technology to detect their eye movements. The stimuli in these studies seem to more often assess understanding of social events rather than physical events, and they also typically measure anticipatory or preferential looking, rather than comparing total time looking at the display across different events, as in VOE studies (see Hopper et al., 2021; Lewis & Krupenye, 2022 for reviews). This is potentially because social events are more likely to hold an ape's attention than are physical events, or because depicting physical events on a 2-dimensional display both requires greater mental representation (as individuals must reason about the 2-dimensional display as a representation of 3-dimensional space) and because it potentially lacks ecological validity (as the scene on the screen may feel particularly removed, and therefore less relevant, to the ape).

3.1.2 Primates' core knowledge of objects: Evidence from search tasks

Numerous studies have used primates' searching behavior as a measure of their expectations about core knowledge principles of objects (Cacchione et al., 2009; Hauser, 2001; Hauser et al., 2001; Hood et al., 1999; Santos, 2004; Tomonaga et al., 2007). These studies typically require individuals to employ their understanding of principles of objects and their motion in order to locate hidden food. In many of these studies, experimenters showed the primate participant a food item being dropped either behind an occluder or into a tube and then allowed the participant to search for the food. This seems to be a difficult problem for primates to solve, and individuals often search in the location directly beneath where the food was released, seemingly failing to integrate information about the solid shelves or tube walls that would impact the food's location. This phenomena has been observed in cotton-top tamarins (*Saguinus oedipus*) (Hauser et al., 2001; Hood et al., 1999), rhesus macaques (Hauser, 2001), chimpanzees (Cacchione et al., 2009;

Tomonaga et al., 2007), and orangutans (*Pongo pygmaeus*), gorillas (*Gorilla gorilla*), and bonobos (Cacchione et al., 2009). Results are somewhat more mixed with similar horizontal displacements, which require less reasoning about gravity but still requires an understanding of solidity and an ability to reason about the spatiotemporal dynamics of a hidden object. Cotton-top tamarins perform somewhat better on the horizontal version of the task than the vertical version, but they still struggle to reliably locate the food (Hauser et al., 2001). Rhesus macaques have demonstrated mixed competency, succeeding in one study (Hauser, 2001) but failing in another (Santos, 2004). Great apes seem to be more proficient at this task, as all four species reliably chose the correct location (Cacchione et al., 2009). Interestingly, Santos (2004) found that rhesus macaques succeeded on the horizontal version of the task when more of the display was visible, such that the possible final positions of the objects remained occluded but the rest of the display was not, providing additional spatiotemporal information. She suggests that this indicates that the monkeys were relying on spatiotemporal information to guide their reasoning and struggled to mentally represent the contact-mechanical dynamics that occurred behind occlusion.

Study 3 contributes to this literature by presenting a study in which apes had to flexibly deploy different aspects of core knowledge of objects in order to locate hidden food items. We aim to gain a more holistic understanding of how humans' closest living relatives reason about core knowledge of objects in order to gain greater insight into our potentially shared foundational understanding of objects and how they interact with their environments. Past studies of this nature with apes have focused on a single facet of core knowledge at a time, making it more likely that apes are able to pick up on surface features of the task over repeated trials and may not reflect their ability to reason flexibly about different aspects of objects and their motion. In additional, many of the studies in which primates failed to reliably locate the food required participants to track the motion of an object once it is released and begins to fall; the present study instead tests the ability to reason about object support if the object is initially stationary.

3.2 Study 3: Bonobos use principles of core knowledge of physics to reason about hidden food

In Study 3, we assessed whether bonobos could successfully reason about the location of hidden food based on the principles of (1) solidity, that one solid object cannot pass through another solid object, (2) support, that an object cannot hover in midair, and (3) spatiotemporal continuity, that objects are bounded entities that can only exist in one place at any given time. We designed a paradigm, in which apes watched as grapes were hidden in one location of a box apparatus and a manipulation was then performed that required apes to deploy one of these principles of core knowledge to correctly locate the grapes. This paradigm was initially inspired by one designed by Feigenson et al. (2002) to test number competency, in which food items were sequentially placed from a central container into two choice locations and infants could then crawl to the location of their choice, typically the side that had more food.

3.2.1 Methods

Participants

All data was collected at the Ape Cognition and Conservation Initiative (ACCI). ACCI is a 13,000 square foot facility designed to support noninvasive research with bonobos. The facility consists of a number of indoor and outdoor (approximately 6-acre ape-accessible yard) housing and research areas. This sample included 5 adolescent and adult bonobos that presently reside at ACCI (1 female, mean age = 21 years, age range = 10 years to 40.7 years). Two additional apes attempted several trials but did not wish to engage with the task and were therefore excluded from this study. All apes completed 76 trials; if a trial was aborted in real time or based on video review, that trial was rerun at the end of the study. Exclusion criteria were determined prior to data collection.

All apes have been born and raised in captivity and have had extensive experience interacting with and receiving food from humans. Furthermore, all apes had previous experience participating in tasks that required them to use pointing to make a choice and therefore required no additional training for this study. All experiments at ACCI are carried out on a voluntary basis, as the apes are able to communicate whether they want to participate on a given day and are always free to step away from the experiment if they wish. Their food intake is not dependent on their performance or participation in experiments; they get the same amount of food no matter whether they participate in experiments and no matter their performance in the study. This study was approved by the ACCI IACUC (protocol #190904-04) and adhered to all guidelines as such.

Design

There were three main trial types, which aimed to assess the apes' ability to reason using (1) spatiotemporal continuity, (2) solidity, and (3) support. There were also two verification trial types to ensure that the apes were using the information provided to them in making their selection, rather than relying on a side bias or choosing at random. These two verification trial types involved choosing between (1) fully visible options, or (2) fully visible then fully occluded options. All conditions were counterbalanced for side and order (when applicable).

Spatiotemporal continuity, solidity, and support conditions aimed to test whether bonobos could reason about the location of the grapes using the principles that (1) a given object can only exist in one space at a time, (2) one solid object cannot pass through another solid object, and (3) an object cannot hover in midair, but must instead rest upon some surface, respectively. In spatiotemporal continuity trials, the experimenter repeatedly reached between a cup of grapes beneath a table and two boxes atop the table. If apes use spatiotemporal continuity to reason about the location of the grapes, they should recognize that if the experimenter’s hand comes out of one of the boxes empty and then reaches down to the cup and reappears holding a grape, it must therefore be an additional grape because the grape that was placed in the box could not have moved down to the cup on its own. In solidity trials, the experimenter rolls grapes from one side of a box to another, either with or without a dividing wall in the middle of the box. If apes reason via solidity, they should recognize that the grapes will pass through to the other side of the box when there is not a wall but will stay on the same side of the box if there is a wall because the grapes cannot pass through. In the support condition, the experimenter places grapes on a shelf and then either removes that shelf or another shelf. If apes reason about the support dynamics of the shelf, they should expect that the grapes will fall when that shelf is removed but will remain stationary if an irrelevant shelf is removed.

Each ape completed 5 trials of each test condition and 2 trials of each verification trial type, for a total of 76 trials per bonobo (20 solidity, 20 support, 20 spatiotemporal continuity, 16 verification trials). Because this study sampled a small number of participants, each with many repeated trials, power analysis (whether via simulation or canned routine) was challenging and likely uninformative because it relies heavily on an a priori estimate of the degree of interdependence between trials. However, the extent of this interdependence was not known before data were collected and therefore a wide range of possible interdependence (correlation) values would have to be entered into the power calculation. In addition, as this is a new paradigm that has never before been used with non-human primates, the effect sizes that would go into the power simulation would be largely guesswork, necessitating that a range of effect sizes be used. Both of these issues would have resulted in an extremely wide range of trial number estimates for the study, which would not have been particularly informative. Therefore, we decided to run 20 of each trial type per bonobo. We believe that this balanced getting as much information as possible about performance on the different trial types with the limitations on the apes' patience and willingness to sit through repeated trials and also to limit to the extent to which the apes could learn from the surface features of the task.

Procedure

Bonobos sat facing the experimenter, who knelt behind a table. The tabletop was on tracks that allowed it to slide forward towards the bonobo when it was time to prompt the participant to make a choice. Bonobos only participated when they were alone in a room. The experimenter and the apes were seated only several feet apart and had clear visual and auditory access to one another. Music was played throughout testing so as to minimize auditory cues regarding the location of the grapes.

Before each session, the experimenter placed an object on a piece of posterboard and tilted it back and forth to demonstrate how an object slides when tilted. During each session, bonobos were shown up to 6 trials. All trial types were interspersed and were run in a randomized order.

For each trial, the experimenter first placed the appropriate apparatus on the table. She then showed the ape the relevant manipulation (see below). The experimenter then slid table forward and said, "which one?" with her head down, so as not to prompt the ape. Once the ape pointed clearly to one box over the other, the experimenter pulled the table back, opened the chosen location and gave the grapes to the ape, if there were any at that location. She then opened the other location(s) to reveal the grapes inside (or that it was empty) before removing them and returning them to the cup underneath the table, if there were any. If the ape did not point when the table was slid forward, or if their point was ambiguous, the experimenter slid the table back and then pushed it forward toward the ape again, and again said, “Which one?”

Spatiotemporal continuity trials began with two boxes with removable fronts resting on the sliding table (Figure 3.1). There was a large box underneath the table which hid a container of grapes from the ape’s view. The experimenter first opened the fronts of both boxes simultaneously, drew the ape’s attention to demonstrate that both were empty, then closed both boxes simultaneously. For part A of these trials, the experimenter reached down and picked up one grape, waved it over box A, placed it into the box, and then waved her empty hand over the box. She repeated this three more times, leaving four grapes total in the box. For part B of these trials, the experimenter reached down and picked up one grape, waved it over box B, reached into the box, pulled her hand back out still holding the grape and waved it over the box, then reached back down to the grape cup under the table. She repeated this two more times. The fourth and final time that she reached into the box with the grape, she placed the grape in the box, and then waved her empty hand over the box. Therefore, at the end of the manipulation, one box had four grapes inside and the other had one grape. Side and order of Part A and Part B were counterbalanced.

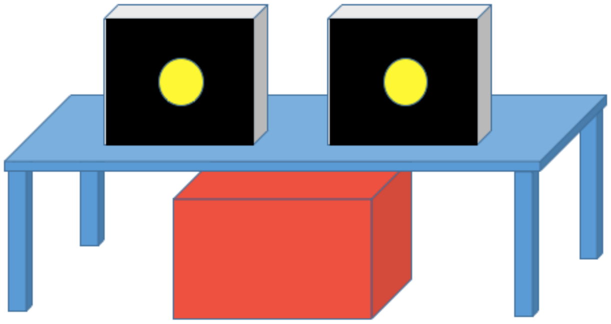

Figure 3.1 Apparatus for spatiotemporal continuity trials

Solidity trials began with two conjoined boxes (either with or without walls on the joining sides, depending on condition) resting on a sliding table (Figures 3.2 and 3.3). The experimenter separated the two boxes and turned them towards the ape to show the existence or lack of the fourth wall. She then stuck the two boxes together, opened the fronts of both boxes simultaneously to show that the boxes were empty, and either A) waved her hand between the boxes to emphasize that there was no wall separating them, or B) tapped on each side of the dividing wall to emphasize that it separated the two boxes. The experimenter then closed the fronts of both boxes simultaneously. She then reached down, picked up three grapes in one hand, and waved them over either the left or right side so that all three grapes were visible to the ape. She placed the grapes on that side, waved her now empty hand above that side, and then picked up the boxes and tipped it to the opposite side (eg: if grapes were placed on the right, she tipped to the left) and placed it back on the table.

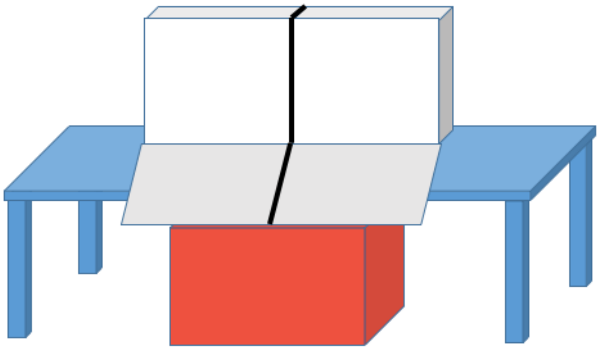

Figure 3.2 Experimental set-up for solidity wall trials

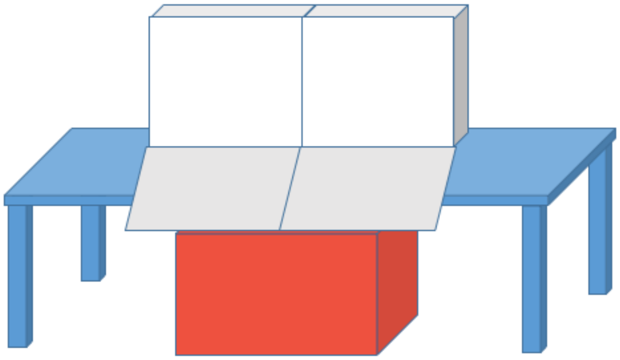

Figure 3.3 Experimental set-up for solidity no-wall trials

Support trials began with a tall box with three doors, one on top of the other, resting on sliding table (Figure 3.4). The experimenter opened all three doors to show that the box was empty. She then slid in two horizontal shelves, which separated the box into three compartments aligned with the three doors. She then reached down, picked up three grapes in one hand and waved them in front of the box so that all 3 grapes were visible to the ape. She placed the grapes on the middle shelf, waved her empty hand, and closed all three doors simultaneously. The experimenter then removed either the top shelf (so the grapes remained on the middle shelf) or the middle shelf (so the grapes fell to the bottom of the box).

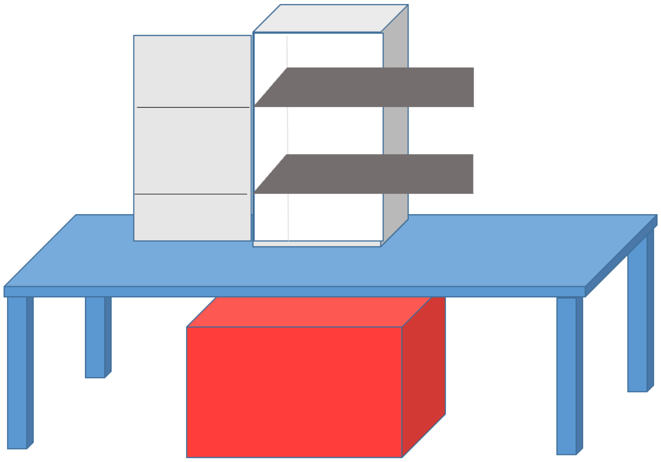

Figure 3.4 Experimental set-up for support trials

Coding and analyses

The bonobos' choices were evaluated in real time because the experimenter needed to reveal the location to which the bonobo pointed and give the bonobos the grapes at that location, if there were any. A reliability coder then coded all videos after data collection was completed. Videos were trimmed to include only the choice by the ape, excluding the manipulation of the apparatus, which side the experimenter thought the ape chose in real time, and as much information about the correct choice as possible (though for some trials this was unavoidable, depending on the apparatus used and the angle of the camera). The intraclass correlation (ICC) between the live coder and the video coder was 0.832, 95% confidence interval [0.789, 0.861]. Given this high degree of agreement, the secondary coder's codes were used for analyses, as she coded with less contextual information.

Choices were converted to a binomial, correct (1) or incorrect (0). A choice was considered correct when the ape chose the location where there were more grapes. Verification trials were removed from the data set and were analyzed separately as exclusion criteria, with a threshold of 50%, such that if bonobos chose incorrectly on more than half of verification trials, they would be excluded from the final sample. Due to issues with model convergence with a frequentist approach, we used Bernoulli Bayesian regression models for each event type in R (version 4.2.1; RCore Team, 2013), using the brms package (Bürkner, 2021) and the tidybayes package (Kay, 2022). We used Markov chain Monte Carlo (mcmc) sampling to assess the posterior distribution space of each model. Because participants were tested across multiple trials, all models included subject identity as a random effect in order to account for repeated measures. We conducted two categories of analyses for this data, (1) models with an intercept only to assess performance for each event type compared to chance, and (2) hypothesis-driven models preregistered on Open Science Framework[3]. For all analyses, we used default priors, 2000 iterations and a warmup period of 500 iterations. The models assessing performance used 2 chains and the models assessing factors that impacted performance used 4 chains. For the first of these two sets of models, we present the intercept, indicating the bonobos' performance on each task, along with the proportion of the posterior distribution that is above chance. For the second of these two sets of models, we used the Region of Practical Equivalence (ROPE) method using the BayestestR package, which assesses the percentage of the posterior distribution that is within a pre-defined null region. We used the default null region of -0.1 to 0.1, and values less than

---

[3] The model presented here differed in several ways from the preregistered analysis, due to the structure of the tasks and issues that arose during data collection. Specifically, (1) models were run separately for each event type, rather than running a single model with event type as a fixed effect, (2) the support model did not include side of the correct answer as a fixed effect or a random slope because the correct location was inherent to the condition (rather than being a feature of counterbalancing), (3) trial number was excluded due to collinearity with trial number of a given condition, and (4) sex was removed as a factor because only one female participated.

2.5% are considered to be evidence of the alternative hypothesis (Makowski et al., 2019). In addition, we present raw values for individual bonobo performance to reflect individual variation in performance on each task.

3.2.2 Results

All 5 bonobos in the sample performed above chance on verification trials and were therefore included in the final sample (Table 3.1). Each hypothesis driven model included age (averaged across the span of data collection) and trial number of that condition (to assess learning or loss of attention/motivation over the span of data collection), both of which were scaled, as fixed effects. Additional factors varied by event type.

Table 3.1 Raw proportion of verification trials correct by bonobo, evaluated as inclusion criteria. "Visible" denotes that the two options remained fully visible throughout the trial, and "occluded" denotes that the two baited quantities were covered before the bonobo was prompted to make a choice.

		Kanzi	**Maisha**	**Mali**	**Nyota**	**Teco**
Verification trial performance	Visible	1.00	0.875	0.857	1.00	1.00
	Occluded	1.00	1.00	0.750	1.00	1.00
	Overall	**1.00**	**.938**	**0.800**	**1.00**	**1.00**

Spatiotemporal continuity

According to our model that included only an intercept and a random effect of bonobo identity, we found that the bonobos' probability of success on the spatiotemporal continuity trials overall was .619, with 91.50% of the posterior distribution above chance (Figure 3.5).

To assess the factors that affected performance, we ran an additional hypothesis-driven model. In addition to age and trial number of each condition, there was an additional fixed effect of side of the correct response. We also included correct response grouped by bonobo identity as a random slope to account for side biases that might differ between bonobos. This analysis

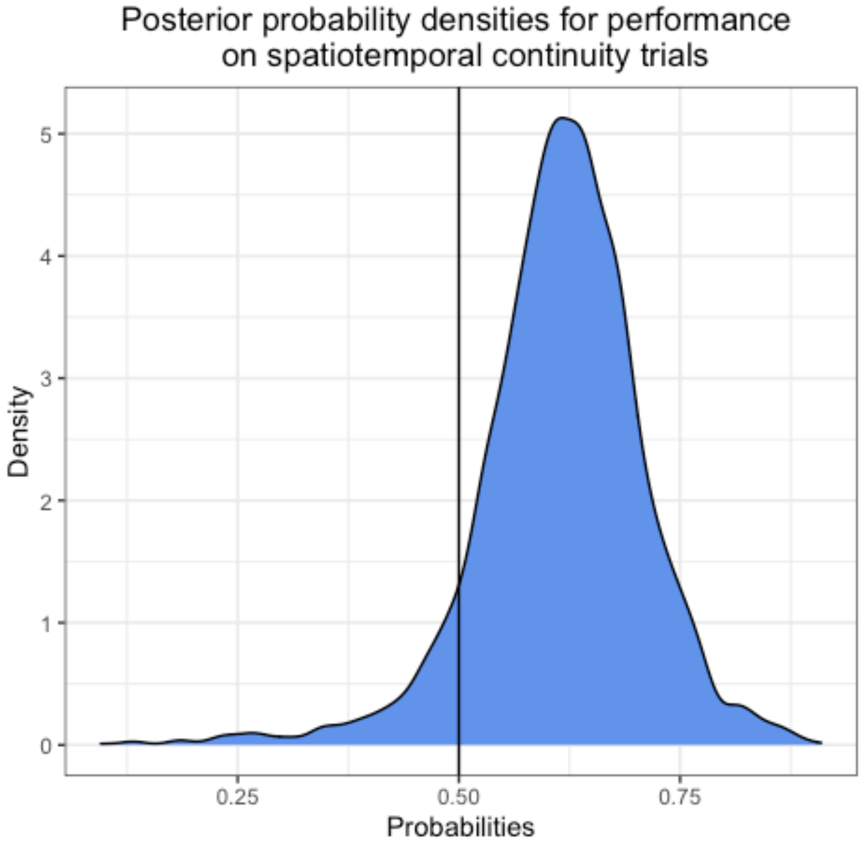

Figure 3.5 Density plot for the posterior distribution of the model that included only an intercept and bonobo identity as a random effect to assess performance compared to chance. The black line represents chance.

revealed that performance on spatiotemporal continuity trials was not affected by the trial number of the condition (all spatiotemporal continuity tests, as there were no distinct condition types, only counterbalancing) (ROPE = 96.90%, Est = -0.02, credible intervals (CI) [-0.10, 0.07]), age (ROPE = 92.53%, Est = 0.03, 95% CI [-0.06, 0.13]), or side of the correct answer (ROPE = 4.12%, Est = 0.92, 95% CI [-1.08, 2.85]). An exploratory model included which quantity (1 or 4) was shown first to test for potential order effects and found no effect (ROPE = 5.97%, Est = -0.68, 95% CI [-1.67, 0.29]). Bonobos had differential performance on the spatiotemporal continuity task, with raw proportion of trials correct ranking as follows: Kanzi (0.750) and Nyota (0.750), Teco (0.611), Mali (0.474), and Maisha (0.471) (Table 3.2).

Table 3.2 Raw performance on spatiotemporal continuity tasks, by bonobo and overall. "4 first" denotes the counterbalancing condition where the first box is baited with four grapes and the second box is baited with one grape. "1 first" denotes the reverse.

		Kanzi	**Maisha**	**Mali**	**Nyota**	**Teco**	**Mean**	**SD**
STC	4 first	.600	.500	.333	.800	.556	.558	.169
	1 first	.900	.444	.600	.700	.667	.662	.165
	Overall	**.750**	**.471**	**.474**	**.750**	**.611**	**.611**	**.139**

Solidity

According to our model that included only an intercept and a random effect of bonobo identity, we found that the bonobos' probability of success on the solidity trials overall was .601, with 91.53% of the posterior distribution above chance (Figure 3.6).

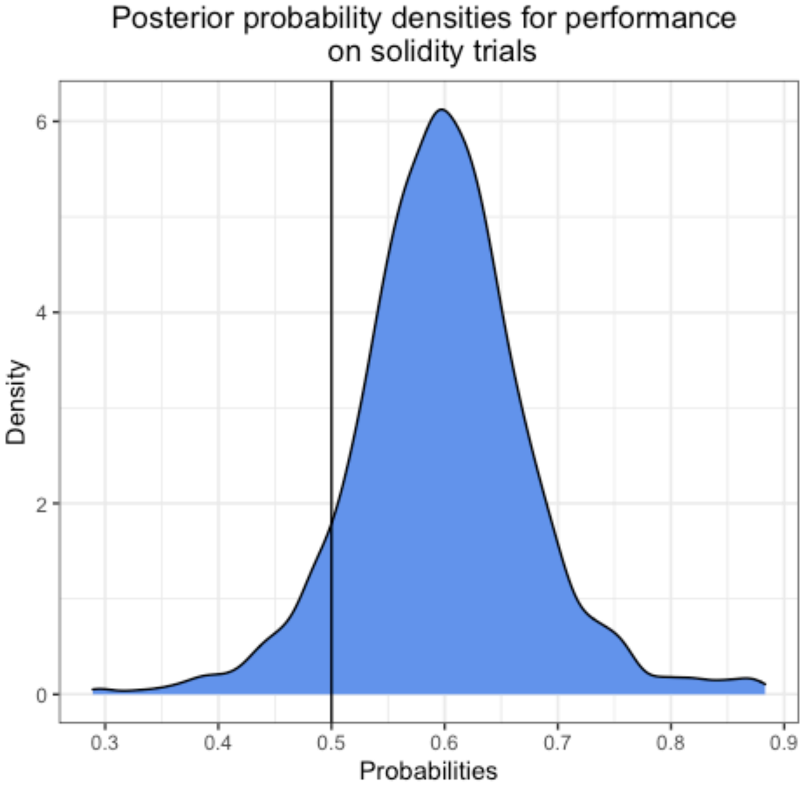

Figure 3.6 Density plot for the posterior distribution of the model that included only an intercept and bonobo identity as a random effect to assess performance compared to chance. The black line represents chance.

In addition to age and trial number of each condition, fixed effects in the hypothesis-driven model included condition (wall vs. no wall) and side of the correct response. We also included correct response grouped by bonobo identity as a random slope to account for side biases that might differ between bonobos. This analysis revealed that performance on solidity trials was not affected by condition (ROPE = 14.20%, Est = -0.21, 95% CI [-1.20, 0.80]), trial number of the condition (solidity- wall or solidity – no wall) (ROPE = 49.53%, Est = -0.10, 95% CI [-0.29, 0.08]), age (ROPE = 82.65%, Est = 0.05, 85% CI [-0.06, 0.20]), or side of the correct answer (ROPE = 2.65%, Est = 1.54, 95% CI [-2.82, 5.77]). Bonobos had differential performance on the solidity task, with raw proportion of trials correct ranking as follows: Kanzi (0.700), Nyota (0.650), Teco (0.632), Mali (0.500), and Maisha (0.474) (Table 3.3).

Table 3.3 Raw performance on solidity trials, by bonobo and overall. “No wall” denotes the condition where the two boxes become a single box when attached. “With wall” denotes the condition where there is a wall that separates the two boxes when attached.

		Kanzi	**Maisha**	**Mali**	**Nyota**	**Teco**	**Mean**	**SD**
Solidity	No wall	.700	.500	.500	.600	.778	.616	.123
	With wall	.700	.444	.500	.700	.500	.569	.122
	Overall	**.700**	**.474**	**.500**	**.650**	**.632**	**.591**	**.099**

Support

According to our model that included only an intercept and a random effect of bonobo identity, we found that the bonobos’ probability of success on the solidity trials overall was .520, with 61.6% of the posterior distribution above chance (Figure 3.7). Note that while chance in this task was technically 0.333, as there were three doors, no bonobo ever chose the top location. We therefore considered only the middle and bottom door, yielding a chance performance of 0.5 when choosing between these two locations.

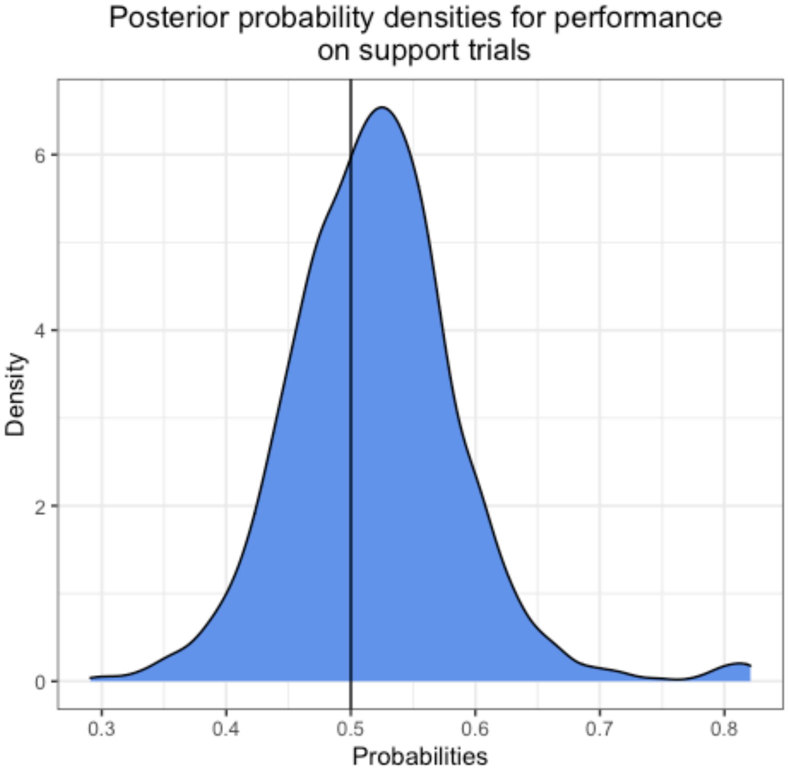

Figure 3.7 Density plot for the posterior distribution of the model that included only an intercept and bonobo identity as a random effect to assess performance compared to chance. The black line represents chance.

In addition to age and trial number of each condition (or the middle or bottom shelf condition), there was an additional fixed effect of condition (middle shelf or bottom shelf removed). This analysis revealed that performance on support trials was not affected by condition (ROPE = 10.305, Est = -0.47, 95% CI [-1.29, 0.35]), trial number of the condition (ROPE = 81.72%, Est = -0.02, 95% CI [-0.16, 0.13]), or age (ROPE = 98.32%, Est = 0.01, 95% CI [-0.05, 0.08]). Bonobos had differential performance on the support task, with raw proportion of trials correct ranking as follows: Nyota (0.579), Mali, (0.550), Kanzi (0.526), Maisha (0.500), and Teco (0.429). Bonobos tended to show highly differential performance across the two manipulations, with Kanzi, Mali, and Maisha getting 100% of trials right in one condition but only 0% to 10% correct in the other. Similarly, Teco had 80% correct when the top shelf was removed and only 9% when the middle shelf was removed, and Nyota had 44% correct when the top shelf was removed and 70% when the middle shelf was removed (Table 3.4).

Table 3.4 Raw performance on support trials, by bonobo and overall. "Top shelf removed" denotes the condition where the top shelf is shelf that is removed, such that the grapes remain in the baited location. "Middle shelf removed" denotes the condition where the middle shelf is removed, such that the grapes fall to the bottom location.

		Kanzi	**Maisha**	**Mali**	**Nyota**	**Teco**	**Mean**	**SD**
Support	Top shelf removed	1.00	.000	.100	.444	.800	.469	.433
	Middle shelf removed	0.100	1.00	1.00	.700	.091	.578	.457
	Overall	**.526**	**.500**	**.550**	**.579**	**.429**	**.518**	**.057**

Table 3.5 Combination of Tables 3.2-3.4 for ease of comparison

		Kanzi	**Maisha**	**Mali**	**Nyota**	**Teco**	**Mean**	**SD**
STC	4 first	.600	.500	.333	.800	.556	.558	.169
	1 first	.900	.444	.600	.700	.667	.662	.165
	Overall	**.750**	**.471**	**.474**	**.750**	**.611**	**.611**	**.139**
Solidity	No wall	.700	.500	.500	.600	.778	.616	.123
	With wall	.700	.444	.500	.700	.500	.569	.122
	Overall	**.700**	**.474**	**.500**	**.650**	**.632**	**.591**	**.099**
Support	Top shelf removed	1.00	.000	.100	.444	.800	.469	.433
	Middle shelf removed	0.100	1.00	1.00	.700	.091	.578	.457
	Overall	**.526**	**.500**	**.550**	**.579**	**.429**	**.518**	**.057**

3.2.3 Discussion

Spatiotemporal continuity

This study revealed that bonobos seem to be able to utilize core knowledge of spatiotemporal continuity to reason about the location of hidden food, as they more often chose the box that had been baited with four grapes as opposed to the box that had been baited with only one grape. In order to accurately track these quantities, apes needed to recognize that when the experimenter reached into the box with a grape and removed her hand without the grape, the grape remained inside the box, and that when she subsequently reached below the table and then showed a grape in her hand again, it must therefore be a different grape. Thus, when she reached into the box again and her hand again reappeared empty, the second grape was added to the quantity in that box. This is in contrast to the manipulation in which the experimenter is still holding the grape when her hand comes back out of the box. In this case, the grape cannot be in the box because it can only be in a single location at a given time, and it is in the experimenter's hand. Thus, there are no grapes in the box until the experimenter finally lets go at the end. The surface features of these two manipulations are identical, as they are matched in movements and both end with the experimenter revealing an empty hand before prompting the ape to choose. Therefore, their choice must rely on their ability to track the location of the grape, using information related to spatiotemporal continuity.

Our hypothesis-driven model revealed that age did not account for any variability in performance, yielding no evidence of developmental change. Importantly, all individuals in our sample were adolescents or adults, so we cannot provide insight into earlier development of this ability. There was also no apparent effect of how many trials each ape had done in this condition, indicating that apes were not learning over the course of the task, nor were they waning in attention or motivation to the point that it impacted performance. Finally, there was no effect of side of the correct answer, such that a side bias would overshadow underlying abilities. Our exploratory model revealed no significant effect of the order of the manipulation. Thus, apes were no more likely to choose the box with four grapes if that box was baited first or second. This indicates that they were able to hold the representation of the hidden quantity in mind as the second box was baited, and they were not just choosing the box most recently baited.

There was a noticeable gap in performance across apes, with Kanzi and Nyota far outperforming Mali and Maisha, and with Teco intermediate to the two clusters. There are a number of potential explanations for this disparity. For one, Kanzi and Nyota were both involved in language-training from a very young age. Therefore, there could be some effect of language on their performance in this task, or it could be that vast experience with participating in experiments of this kind have better attuned them to the subtleties of the manipulations or to tracking the two quantities. At present, we cannot disentangle these two possibilities, but future research should investigate the role of language in reasoning on this task. Additionally, Mali and Maisha showed relatively low levels of attention during this task and needed frequent reorienting. Spatiotemporal continuity trials were particularly long, and bonobos needed to attend as the experimenter reached into each box four times, which was taxing for some apes. In addition to personality differences that may account for differential attention, Mali's performance may also have been impacted by the fact that 1) she was relatively new to the institute and was new to separating from other individuals for research, and 2) as a female, she is more central to group social dynamics, potentially making it more distracting to be separated from the group. Evidence for this second explanation comes from the fact that the other two females in the group were unwilling to separate from the group for the duration of the task, and could not complete trials due to inattention, as they oriented their attention towards the location of groupmates. Despite this individual variation, overall performance on the task indicates that the apes as a whole performed above chance on this task with a greater likelihood to choose the higher quantity. This demonstrates at the very least an ability to utilize the principle of spatiotemporal continuity to reason about the location of hidden food items.

Solidity

This study indicates that bonobos are able to use the core knowledge principle of object solidity to reason about the location of hidden food items. When two boxes are attached to each other such that they form a single box, the bonobos seem to recognize that the grapes will pass through, as they are more likely to choose the side opposite that which was baited. When the boxes have a wall on their connecting sides, however, the bonobos are more likely to choose the same side as where the grapes were placed, indicating that they recognize that the grapes would not be able to pass through this dividing wall. The surface features of these two manipulations were identical, as the boxes looked exactly the same except for the dividing wall, and the experimenter performed all the same motions. Therefore, the bonobos' success on these two manipulations demonstrates their ability to reason flexibly about solidity to track the location of the grapes. These findings are consistent with those of Cacchione et al. (2009), which found that great apes succeeded on task that similarly required individuals to reason about how an object would move along a horizontal trajectory with something blocking the track.

Our hypothesis-driven model revealed no difference in performance between these two conditions, indicating that the apes were attending to whether or not the wall was present and were integrating the presence or absence of the wall into their representation of how the grapes would move when the box was tipped. The model also revealed no effect of age, though as previously mentioned, it is difficult to draw conclusions about the developmental trajectory of this ability without testing younger individuals. However, in our sample, there does not appear to be any change in the ability to reason about solidity between adolescence at older adulthood. There was also no effect of trial number of each condition, which implies that apes were neither improving in performance over the course of the task as they learned, nor were they degrading in performance, as we might expect if attention or motivation were decreasing over repeated trials. Finally, there was no effect of side of the correct answer, so side biases were not overshadowing performance on the task.

In fact, there were noticeable differences in performance on this task across bonobos, with a pattern that very closely matched that of the spatiotemporal continuity task. Kanzi and Nyota demonstrated relatively strong performance across both conditions. As previously described, this could be attributed either to some benefit of language or on their extensive experience with performing experimental tasks. Future research should address these two possibilities. Teco performed better on the condition where the wall was absent as compared to when the wall separated the two boxes, in which case he was at chance performance. It therefore may have been more difficult for him to reason about the fact that the wall would block the grapes' trajectory. Maisha and Mali performed relatively poorly, seemingly choosing randomly between the two options in both conditions. Potential explanations for this disparity mirror those described in the previous section.

While there was individual variation in performance, as a group, the apes tended to more often choose the correct location across solidity trials, indicating that they recognize that the grapes will move differently depending on whether or not a wall divides the box, and that the grapes cannot pass through that solid wall.

Support

The bonobos seemed unable to reason about how grapes would move in vertical space using the core knowledge principle of support. Interestingly, as revealed by our hypothesis-driven model, apes seemed as likely (or unlikely) to choose the correct location of the grapes when the top shelf was removed, which should not have impacted the location of the grapes, as compared to when the middle shelf was removed, which removed support from the grapes, causing them to fall down. This failure is consistent with past studies in which primates had to reason about the vertical movements of objects (Cacchione et al., 2009; Hauser, 2001; Hauser et al., 2001; Hood et al., 1999; Tomonaga et al., 2007). In fact, even infants seem to find it more difficult to reason about object motion when occluded objects are moving vertically rather than horizontally (Hood et al., 2000). This could be because integrating information about how gravity effects object motion interferes with core knowledge principles of object movement, such as support. We attempted to mitigate this effect by having the grapes start out in a stationary position, rather than having the apes watch as the grapes were released and began to fall, but this does not appear to have been an effective solution. In the present task, apes never chose the top location. This could mean that they did not consider this to be one of the options from which they could choose (though the top door was marked in the same way as were the middle or lower doors) or that they did not think the grapes would move up from the middle location.

Patterns of individual performance on the present task indicate that each ape tended to choose a single location across all trials of support events, rather than reasoning flexibly across trials. Four out of five bonobos got all or most of the trials correct in one condition and a very low proportion correct in the other, which indicates that they were always choosing the same location (which was correct when one of the shelves was removed but incorrect when the other shelf was removed). Only one bonobo, Nyota, seemed to change his choice flexibly across support trials, though he chose the bottom location nearly twice as often as the middle location. Neither age nor trial number of each condition significantly impacted performance. This makes sense, as the bonobos' strategy of repeatedly choosing the same location did not seem to vary between individuals of different ages nor across the span of the task.

While the bonobos in this task failed to reason flexibly about how the removal of a supporting surface would impact the movement of hidden food, this by no means provides conclusive evidence that bonobos do not reason about object support. Keen (2003) reviews evidence that human children struggle to choose the correct location in similar search-based tasks that rely on core knowledge of objects, despite the fact that their looking times indicate an underlying understanding of where the object should end up. As she proposes in regard to children, it could be the case that having to *predict* the outcome and act upon that prediction is more difficult than recognizing a violation when it occurs, particularly when the mechanics of the task are complex. Therefore, it is possible that the manipulations in these trials were complex and difficult for the apes to represent, as the shelves were first added to the empty box before grapes were baited, and then were pulled out. Future studies should utilize a simpler display that has fewer moving parts that must be attended to and represented by the apes. In addition, future studies should assess where the apes are looking as the manipulation is conducted, as this could shed light onto their expectations about the objects' movements, even if their executive functioning in the task interferes with them pointing to the location (as in Butler et al., 2002).

In addition to the potential difficulties with the demands of this task, reasoning about the removal of support might be more difficult than recognizing an overall lack of support. Furthermore, this task conflated reasoning about gravity with reasoning about support, which may have impacted performance. Even in looking time tasks, young humans seem to struggle in similar paradigms, looking equally long when a dropped object is revealed to hover as when it appears to rest on a surface (Spelke et al., 1992), and expectations of object support seem to come online more slowly than do other expectations from other principles of core knowledge (Baillargeon, 1995; Baillargeon et al., 1992).

3.3 General discussion

Study 3 provided evidence that bonobos, one of humans two closest living relatives, are able to use principles of core knowledge to reason about objects they cannot see, specifically the principles of spatiotemporal continuity and solidity. This indicates that not only do they recognize a violation when they see one, such as in looking time studies, but they also are able to mentally represent how objects move and interact even when they are out of view. Bonobos in this study were able to track objects and expected them to remain in a given location unless acted upon, in accord with spatiotemporal continuity, and to stop when they hit a wall, in accord with solidity. The apes in our sample seemed unable to reason about how objects would move when support was removed, but future work should aim to test bonobos on a task support task with a simpler design that does not conflate reasoning about gravity so heavily with reasoning about support.

The sample size in this study also limits our ability to make generalizations about bonobos more broadly. While it does indicate an *ability* to reason about spatiotemporal continuity and solidity, it is important to recognize that this is a very small sample, and that individuals at ACCI are enculturated and have extensive experience engaging in experimental tasks. This study should be replicated with a larger sample to gain greater insight into the generalizability of these abilities in bonobos. Future studies should also aim to conduct this task across with younger individuals, to yield insight into the developmental trajectory of these abilities, and with a broader range of taxa, to inform the evolutionary history of these abilities.

Ultimately, we provide evidence here that bonobos, like humans, mentally represent hidden objects and expect them to move in ways that are consistent with principles of core knowledge. Furthermore, they can use this understanding to guide their own actions, in this case, searching for food.

Chapter 4 Surprise-induced exploration in bonobos

4.1 Background

Many adults can relate to the experience of watching a magic trick and thinking to themselves, “wait, how did they do that?”. In fact, given the opportunity, many people would probably want to have a look inside the hat that just made a rabbit seemingly disappear or to try for themselves to disentangle the metal rings that seemingly passed through one another to link together. This motivation to explore something that violates one’s expectations is not unique to people faced with a magician and her bag of tricks, however. There is evidence that young children, and even infants, prefer to explore something that surprises them. In other words, surprise can serve to motivate a learner to search for an explanation, which in turn can help to reveal an underlying causal mechanism (Charlesworth, 1969; Reisenzein et al., 2019). Perhaps one way that humans achieve their unique levels of abstract causal reasoning is by seeking explanations following surprising events, such that they reveal previously opaque causal mechanisms and are able to integrate these causal relations into their mental representations of how entities interact in the world. Very little research has actually addressed the question of whether this form of surprise-induced exploration is in fact unique to humans. The present research seeks to deepen our understanding of the evolutionary history of this form of explanation-seeking by testing whether one of humans’ two closest living relatives prefer to explore objects that surprise them over those that do not.

4.1.1 Surprise-induced exploration in human infants and children

Perhaps the most foundational evidence that human infants prefer to explore things that surprise them is the fact that infants look longer at scenes that violate their expectations than those that do not (see Stahl, 2015 for this claim and for review). However, once infants are old enough to physically interact with their environments, they engage in more active forms of surprise-induced exploration. In one study, 6-11-month-old infants were presented with a variety of objects, some of which appeared to have a regular flat surface, some of which were concave, and some of which appeared to be concave but had actually been manipulated to be flat. Thus the "surprising" objects in this case were the ones that appeared to be concave but were actually flat. By 9 months of age, infants began to spend more time exploring these surprising objects than either of the other two object types, and this prioritization of the surprising object continued to increase across the age sample. Furthermore, the 9-month-olds, but not younger babies, explored the transparent object for longer stints at a time, and their motions were measured to be slower more deliberative than with either of the other two object types. The authors interpret this as evidence of explanation-seeking behavior (Pieraut-Le Bonniec, 1985).

Stahl & Feigenson (2015) built on this finding by demonstrating that infants not only preferentially explore objects that themselves have some surprising property, but that 11-month-old infants prefer to explore an object that behaves in a surprising way. Specifically, they presented infants with either solidity or support events. Infants in the solidity condition saw an object roll down a ramp, and it was either revealed to have stopped when it hit a wall that was blocking the track (unsurprising) or to have appeared to have rolled through the blocking wall (surprising). Infants in the support condition saw an object pushed across the top of a box, and the object either maintained constant contact with the supporting box (unsurprising) or the object was pushed past the surface of the box but continued to hover (surprising). At test, infants were presented with the toy they had just watched in the display along with a novel object they had never seen. Only infants who had just seen the familiar object do something surprising (roll through the wall or hover) spent more time exploring the familiar object relative to the novel object. Further evidence that this exploration is actually a form of explanation-seeking and not just heightened attention to the surprising object comes from the fact that infants actually tailored their exploration to the type of violation they had just witnessed. Specifically, infants who had just seen the object roll through a wall were more likely to bang the object during the exploration phase, appearing to test the solidity of the object. In contrast, infants who saw the object float in midair were more likely to drop the object, appearing to test the support of the object.

Perez & Feigenson (2022) replicated the surprise-induced exploration effect to solidity violations, and they provided compelling evidence that infants were in fact seeking an explanation through their exploration and were not just trying to replicate the surprising event they had just seen. After infants were shown that the object ended up on the far side of the wall that blocked the track, the wall was turned to reveal that it had a large hole in the bottom, such that an object could pass through. Not only were the infants who were shown this explanation less likely to preferentially explore the familiar object, but at an individual level, the infants who spent the most time looking at the hole in the wall spent the least time preferentially exploring the familiar object. It therefore seems to be the case that infants were integrating the explanation of the hole in the wall into their mental representation of the event. Thus, what was initially surprising was rendered unsurprising and did not require further exploration to reveal an underlying causal explanation.

A final piece of evidence of surprise-induced exploration in human infancy comes from a study in which 13-month-old infants were presented with two populations of balls of varied colors, which were visible to the infant. From one box, an experimenter “randomly” pulled four balls of different colors, a plausible outcome given the population of the box. From a second box, and experimenter "randomly" pulled four balls of the same color, a rather surprising outcome given the ratio of colored balls in the box. Infants were then allowed to approach the boxes. Infants were more likely to crawl towards the surprising box and to search on that side. The preference for the side where the four sampled balls were all the same color did not exist when the balls were pulled from the experimenter's pocket rather than randomly from the box, indicating that it is actually because the sampled distribution was surprising rather than some inherent preference for the pattern. This study provides evidence that infants prefer to explore something that violates their expectations even when those expectations are ones that are obtained over the course of development and are not innate (Sim & Xu, 2017).

Surprise-induced exploration and explanation-seeking of this kind is not limited to infancy. When presented with a wall that falls forwards and back, preschool-aged children were more likely to get up to search when an object was placed behind the wall and it still fell all the way back than when it stopped when it hit the object. Children spontaneously provided verbal explanations for how they thought the violation occurred, and almost every participant searched behind the wall, when given the opportunity, expressing excitement upon discovering the hidden trap-door mechanism (Chandler & Lalonde, 1994). Similarly, Charlesworth (1964) had preschoolers, first-, and third-graders play a game in which they placed marbles into an apparatus. In one condition, the marbles that came out of the apparatus were the same as those that had gone in (unsurprising), but in a second condition the outcoming marbles differed from those input in color and sometimes also in number (surprising). Children in the surprising condition expressed verbal and facial evidence of surprise and wanted to play the game for longer, on average, than did children in the unsurprising condition. While this does not provide direct evidence of children seeking an explanation, it does demonstrate that even in elementary school, children are motivated to explore something more when it violates their expectations.

Children's patterns of surprise-induced exploration map onto their own individual expectations. Bonawitz et al. (2012) took advantage of the natural developmental change in children's theories of balance relationships in order to assess whether children explore more when their conceptions are challenged. Some children believe that an object will balance at its geometric center (typically younger children), while others believe it will balance at its center of mass (typically older children). The participants were then shown identical demonstrations of blocks balancing either at their geometric center or their center of mass. Children explored the block more when it was in conflict with *their* expectations of how an object should balance. Furthermore, children were more likely to posit explanations of an external force (such as a magnet) when they viewed an expectation-violating event. Finally, children were more likely to update their predictions about object balance in light of surprising evidence, provided that an alternative explanation was not present.

Povinelli & Dunphy-Lelii (2001) provide evidence that children prefer to explore surprising objects when the expectations are ones that are taught over the course of the experiment; they need not be innate expectations or even ones that have been entrenched over the course of development. Three- to five-year-olds were presented with a task in which they were to stand blocks on their ends for a reward. After several familiarization trials, the experimenters secretly switched out one of the blocks with one that had been altered such that it could not stand on its end. They found that children tended to examine this block that had violated their expectations, looking closely at it and feeling it with their hands, and they were more likely to try standing it up again in a new location. 5-year-olds were even more likely than were the younger children to engage in such explanation-seeking behaviors, pointing to the possibility of a development trend.

Several studies have demonstrated not only that children explore more when their expectations are violated, but that they actually explore in ways that help them to isolate variables to better understand the violation. In one study, children made predictions about the relative size of the shadows that different objects would cast, when objects were different sizes and were different distances from the display. Children where then shown evidence that was either in accord with their expectations or were in violation of them. They found that all children who witnessed a surprising demonstration subsequently sought to isolate the variables at play, manipulating either the size of the object or its distance from the display, when given the opportunity to explore the objects and the display. On the other hand, only half of children who had not had their expectations violated did so (van Schijndel et al., 2015). Legare (2012) similarly provided evidence of hypothesis testing in situations in which children's expectations are violated. Children were taught that a "blicket detector" would light up only when a "blicket" was placed on top, and they were shown that a certain category of items were blickets. Participants then watched as the box lit up when an object was placed on top, and that object was either a blicket (unsurprising) or not a blicket (surprising). Only when the children saw a non-blicket cause the box to light up did they tailor their exploration to isolate potential variables, placing different combinations of objects atop the box. Furthermore, children were more likely to evoke causal explanations when their expectations were violated (Legare, 2012; Legare et al., 2010).

Taken together, these results demonstrate that young humans, from as early as eight months of age, engage in greater exploration when their expectations are violated. This holds true when expectations are those rooted in innate knowledge, in knowledge or theories that are acquired over the course of development or are presented within the span of the experimental task. Children dedicate more time to exploring the targets of these unexpected events, and they tailor their exploration in ways that would provide themselves with explanations for the events, isolating potential variables and testing potential hypotheses. Thus, this surprise-induced exploration really could be helping young learners to reveal underlying causal mechanisms that might not otherwise be readily available. They can then incorporate this new understanding into their mental model of how the world works and can gain a deeper understanding of object relations that allow them to reason more flexibly about the causal dynamics around them. But are humans unique in their propensity to explore what surprises them?

4.1.2 Surprise-induced exploration in non-humans

To our knowledge, there are only two studies that directly assess whether individuals of a non-human species prefer to explore the target of a surprising event, one in parrots (*Cyanoramphus novaezelandiae*) and the other in chimpanzees (*Pan troglodytes*). The need for such work has been previously emphasized, (eg: Völter et al., 2020), but empirical work on explanation-seeking in non-humans remains remarkably sparse.

In one study, parrots were familiarized to a novel red ball and had ample opportunity to explore the object over the course of several days. The ball was then replaced either with a control object that was identical to the original object, or a test object that differed in one of several properties, either more surface-level features, including color or shape, and or an internal property, its center of gravity. While the birds spent more time exploring any of these altered objects than the control object, they spent the longest time exploring the object that had its center of gravity altered, such that it moved in a surprising way. This demonstrates that it is not just the novelty of the objects that instigated exploration, but actually the change in a property that violated the birds' expectations. They posit that the birds spent more time exploring the object that had an internal property altered in order to gain more information to better understand the surprising event (Demery, 2013). This is in line with past proposals that allocating time to exploring functional cues of an object is more adaptive than exploring surface features of an object because the functional cues are more likely to impact how the object behaves and interacts with its environment and is therefore more relevant for action planning (Chappell et al., 2012).

In the study with chimpanzees, the apes were trained to stand a block on its end and were rewarded when they succeeded (much like in the study with children described in the previous section). In a first experiment, some of the blocks were then altered so that the end was beveled such that the object could no longer stand on its end. In a second experiment, conducted several months later with a new set of blocks, objects were altered internally such that they could not stand on end, but no visible explanation was available. In experimenter 1, chimpanzees did tend to explore the unexpected blocks for longer than the expected ones, and they were somewhat more likely to try to stand the unexpected block in a new location (though this was not statistically significant, $p = 0.06$). Exploration in these cases was primarily exhibited through close visual inspection and sniffing with only a single observation of tactile exploration, in this case mouthing the bottom of the object. To rule out the possibility that preferential exploration of the surprising blocks was due solely to the novelty of the block itself (as the end had been beveled), the altered blocks in experiment 2 had no visible alterations. In this study, only one ape was observed to examine a surprising block. This stands in contrast to children in an analogous study who were significantly more likely to examine the surprising block. There are a number of reasons to interpret these results with caution, however (as discussed in Völter et al., 2020). For one, this task was likely not intuitive or ecologically valid for the chimpanzees, as this experiment required more than six months of training prior to data collection, and placing objects in this way is not a natural behavior. In addition, this extensive training may have focused apes' attention on performing the behavior correctly in order to receive the reward and may have overshadowed motivation to explore. Importantly, apes also were more likely to explore the unexpected block in the first experiment, and only showed a lack of preferential exploration in the second study. Given that these two experiments were conducted on the same population of apes, it is plausible that the unexpected blocks in second experiment were not so unexpected given that they had had prior experience with blocks that failed to stand. Thus, while this study provides a first step towards gaining insight into explanation seeking in apes, it leaves much room for improvement.

Study 4 builds on this literature by assessing surprise-induced exploration in bonobos (*Pan paniscus*). We believe that bonobos can serve as a useful starting point in addressing the possibility of such exploration in a non-human species because 1) they are one of the two species that are most closely related to humans and therefore shed insight into human evolution and human uniqueness, and 2) they are more juvenilized and exhibit higher levels of exploration throughout adulthood than do chimpanzees and other apes (Hare et al., 2012). We designed a non-verbal task that required no prior training or experience and could be run nearly identically with apes and infants, allowing potential for a direct species comparison.

4.2 Study 4: Do bonobos prefer to explore an object that has unexpected properties?

This study seeks to assess whether bonobos, like humans, prefer to explore objects that violate their expectations, in comparison to those that do not. As reviewed above, humans often engage in exploration upon seeing something surprising, and they do so in a way that indicates they are seeking explanations for the surprising events, even before they are a year old. This could potentially help learners to reveal causal mechanisms that might initially have been opaque, and they can then update their mental representation of the event. In this manner, surprise-induced exploration could serve to scaffold abstract causal reasoning. Humans are able to reason about causality more abstractly than are even our closest primate relatives. We posit that surprise-induced exploration and explanation-seeking could be one potential mechanism by which humans achieve their unique flexibility and abstraction in their causal reasoning.

This study was conducted with a group of captive bonobos. We designed a task that required no prior training by capitalizing on the apes' natural tendency to engage with a novel object in their environment. We surprised the apes by dramatically manipulating the weight of objects; we first gave them experience with an object they had never seen before, and we then switched that out for an identical object that was either the same weight (unsurprising) or was much heavier or lighter (surprising). Research with humans demonstrates that even within the course of a single experiment, adults quickly form expectations about an object's weight after minimal contact with the object. Prior expectations based on the object's properties, such as its size or material, impact one's initial interaction with an object, but experience with the object quickly informs one's expectations (Buckingham et al., 2009; Flanagan et al., 2008; Gordon et al., 1993; Johansson & Westling, 1988). These studies demonstrated that humans expect that objects that appear to be identical will have the same properties, including that they will weigh the same as each other. Research has shown that apes, too, form expectations about object weight, and plan their actions depending on the predicted weight on an object (Sirianni et al., 2018). As discussed in the previous section, children preferentially explore objects that surprise them even when the expectations that are violated by that object are ones that were formed over the course of the experiment (Legare, 2012; Povinelli & Dunphy-Lelii, 2001). We therefore believe that it is plausible that apes would form an expectation of an object's weight based on their experience handling that object and would therefore expect an identical object to weigh the same amount. When we change the weight of the object between handling bouts, this should therefore violate the apes' expectations, thus giving us the opportunity to investigate how they differentially explore objects that surprise them.

4.2.1 Methods

Subjects

This sample included 6 adolescent and adult bonobos that presently reside at the Ape Cognition and Conservation Initiative (ACCI) (mean age = 22.7 years, age range = 11.5 to 41.3 years; 2 females). Four apes successfully completed all 8 sessions, one ape completed only 5 sessions (because of refusal to engage with the object during familiarization), and one ape successfully completed 7 sessions (trial was terminated during familiarization because the ape was engaging with the object in a manner that was deemed unsafe by veterinary staff). One additional ape was given numerous opportunities to participate in this study and did not wish to enter the room and/or engage with the objects and thus was not included in the sample.

All apes have been born and raised in captivity and have had extensive experience with man-made objects. As in Study 3, all experiments are carried out on a voluntary basis, as the apes are able to communicate whether they want to participate on a given day and are always free to step away from the experiment if they wish. Their food intake is not dependent on their performance or participation in the experiments. They get the same amount of food no matter whether they participate in experiments and no matter their performance in the study. This study was approved by the ACCI IACUC (protocol #190904-04) and adhered to all guidelines as such.

Design

Bonobos were presented with 3D-printed objects (Figure 4.1). Each object had a heavy version (printed at higher fill) and a light version (printed at lower fill). In each session, bonobos participated in two trials, a familiarization trial and a test trial. For each trial, the bonobo entered the testing room where the object was already waiting, and interactions with that object were observed and recorded. Regardless of whether the object was used for a surprising or unsurprising trial, the object presented to the bonobo during familiarization was distinct from the object presented at test, so as to eliminate any recognition of the object, such as by smell, and

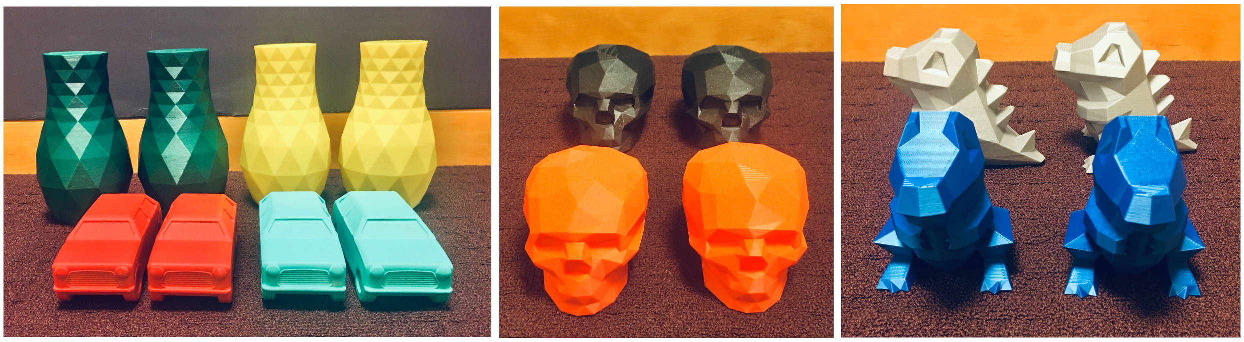

Figure 4.1 Objects used in Study 4, all 3D-printed in polylactic acid (PLA). All templates were sources free from Thingiverse. Green vases, teal cars, black skulls, and silver dinosaurs were used for expected trials, and yellow vases, red cars, orange skulls, and blue dinosaurs were used for unexpected trials.

also because the bonobos often damaged the objects during familiarization, and it would therefore seem surprising if the second object was then presented fully intact.

The two main condition types were whether the weight of the object at test was unsurprising or surprising. In unsurprising trials, the object presented at test was the same weight as the object that was presented during familiarization, and in surprising trials, the test object was either significantly heavier or lighter than the familiarization object (Table 4.1). Each object type (eg: vases) was used for one surprising session and one unsurprising session to account for potential differences in interest to different objects. While the same object type was used for surprising and unsurprising trials, different colored objects were used for the two sessions so that it was not surprising that the object seemed to have changed weights from one session to another.

Table 4.1 Summary of condition design for Study 4. The Familiarization column indicates the weight of the object during the familiarization trial, the Test column indicates the weight of the object during the test trial, and the Condition type trial indicates whether the session was unsurprising (the test object was the same weight as the familiarization object) or surprising (the test object differed in weight from the familiarization object).

Familiarization	Test	Condition type
Light	Light	Unsurprising
Light	Heavy	Surprising
Heavy	Light	Unsurprising
Heavy	Heavy	Surprising

For example, for the light–unsurprising condition, the vases were green, and for the light–surprising condition, the vases were yellow. This is because it was important that the apes were able to clearly discern that the objects in the opposite condition were indeed distinct objects; if a bonobo first participated in a surprising condition, where the object changed weights between trials, the subsequent unsurprising condition might therefore seem surprising because the familiarization object would have reverted back to the weight of the original familiarization object. Order of conditions for each object type (unsurprising vs surprising) was counterbalanced within and between subjects, and order of object (vase vs car; skull vs dinosaur) was counterbalanced between subjects.

Procedure

Bonobos were separated from other apes for each session. Exploration was coded live via CyberTracker® using the ethogram below (Table 4.2) and was also recorded on video (with at least two cameras to maximize visibility) for offline review.

For all trials, the object was in the room upon the bonobo's entry. The trial began once the bonobo touched the object and ended after the bonobo had not engaged with the object for 3 consecutive minutes. The bonobos first participated in a familiarization trial, then waited for 5-10 minutes before progressing into the next room (or out and back into the same room) for the test trial. If for any trial the ape did not interact with the object for 5 minutes, they were moved back out of the room and the object was moved to a new location to encourage engagement. Prior to the start of a test trial, an identical replica of the object from the recent familiarization trial was placed in the room. The second object was visible, though not accessible, to the apes throughout the trials because if a bonobo damaged the object during the familiarization trial, it would then be surprising if the object appeared to be once again intact during the test trial. Apes were also instructed to return the object to a member of the care staff after the familiarization trial to emphasize that the object was not being left in the testing room. Note that care staff were not present during the trials so that the apes were not motivated to return the object for a reward or in order to end the trial.

Participants could only participate in one session per day. Data was collected in two two-week trips, with the goal of 4 sessions per bonobo per trip. Two objects were used during the first trip, vases and cars, and two objects during the second trip, skulls and dinosaurs.

Coding & analyses

We coded exploration by documenting when bonobos were touching the object and looking at the object (while touching it), as well as any behaviors performed on the object. Touching, looking, and dominance displaying were recorded continuously (i.e. their duration was recorded), and all other behaviors were recorded instantaneously (i.e. the behavior was recorded at a single point in time) (see ethogram below for specific behaviors and definitions, Table 4.2).

Table 4.2 Ethogram used to collect behavioral observations of object exploration. Behavior type column indicates whether the behavior was recorded continuously or instantaneously, Behavior column indicates how the behavior was classified, and Definition column indicates a description of each behavior.

Behavior type	**Behavior**	**Definition**
Continuous	Physical contact	Holding object with hands or feet
Continuous	Visual inspection	Visual attention is oriented to object
Continuous	Display	Using object for display (dragging object along ground)
Instantaneous	Bite/mouth	Putting object in mouth (including looking at object in between)
Instantaneous	Sniff	Bringing object to nose
Instantaneous	Stomp	Hitting object with foot against the ground

Instantaneous	Bat	Hitting/slapping object with hand/foot while object is on a surface
Instantaneous	Throw	Holding object then propelling it
Instantaneous	Bang	Forcefully bringing object into contact with a surface
Instantaneous	Shake	Holding object in hand and jiggling it quickly back and forth
Instantaneous	Other	Any other behavior that was performed on the object. The specific behavior was recorded as a note and was later integrated into behavior counts

The experimenter verified live coding by reviewing the videos after data collection was completed. All trials were reviewed for coding accuracy, and times were updated if they deviated more than 2 seconds from the original code. Visual inspection of the object was difficult to accurately assess over video and was prioritized during live coding because of this, so looking was not updated unless the real-time notes indicated to do so, or a code was missing from the real-time data collection. In this review process, we identified three cases in which the real-time coding indicated the start or end of a look without a corresponding start or end code, and the missing code could not be verified on video. These three instances of looking were removed from the data set, as there was no way to assess the duration of looking. Based on contextual information provided by adjacent codes, however, each of these missing looking bouts were relatively short (a maximum of 8, 12, and 16 seconds) and were each removed without consideration of the event type.

We examined exploration using three separate dependent variables: how long bonobos touched the object (tactile exploration), how long bonobos looked at the object while touching it (visual exploration), and the number of distinct behaviors performed on the object. We first assessed whether any observations were outliers by examining whether any values were 1.5 times the interquartile range lower than the first quartile or higher than the third quartile and found no such observations. Therefore, all observations were included in the analyses. Touch and look durations were both log-transformed for analysis[4], as the log-normal distribution provided a better fit than did the normal distribution (log-likelihood for touching = -238.387 as compared to -270.434, log-likelihood for looking = -195.021 as compared to -237.156)[5]. Due to issues with model convergence with a frequentist approach, we used Bayesian regression models in R (version 4.2.1; RCore Team, 2013), using the brms package (Bürkner, 2021) and the tidybayes package (Kay, 2022). We used Markov chain Monte Carlo (mcmc) sampling to assess the posterior distribution space of each model. For all analyses, we used default priors, 5000 iterations and a warmup period of 1000 iterations and 6 chains. For each outcome measure, we conducted three types of models (1) hypothesis-driven models that included condition type (surprising or unsurprising) and session number (to account for potentially reduced surprised across sessions) as fixed effects and bonobo identity as a random effect (to account for repeated trials), (2) exploratory models that included condition type, object type and object weight (to account for potential variation in interest level in different objects or heavier/lighter objects)[6] as fixed effects and bonobo identity as a random effect, and (3) null models that systematically

---

[4] A constant of 1 was added to all look durations prior to log-transformation because bonobos did not look in some sessions and log-transformation cannot be computed on data containing zeros.
[5] Note that plots and descriptive statistics still use raw looking times so that they are easier to interpret.
[6] Note that we excluded session number from this analysis due to collinearity with object type, since object types were presented in the same order.

excluded fixed effects to assess whether inclusion of those factors significantly impacted the model. To assess the effect of each factor, we used the Region of Practical Equivalence (ROPE) method using the BayestestR package, which assesses the percentage of the posterior distribution that is within a pre-defined null region. We used the default null region of -0.1 to 0.1, and values less than 2.5% are considered to be evidence of the alternative hypothesis (Makowski et al., 2019). To compare the hypothesis-driven and exploratory models to their complementary null models, we used Bayes Factors (BF), where a factor of less than 1 is evidence of the null hypothesis, indicating that the factor in question did not significantly influence the dependent variable in question.

4.2.2 Results

Our hypothesis-driven model for the amount of time bonobos spent touching the objects during test sessions (Figure 4.2) revealed no significant difference in time spent touching the object in surprising versus unsurprising trials (ROPE = 7.88%, Est = -0.56, 95% credible intervals (CI) [-1.38, 0.26]) or across sessions as the study progressed (ROPE = 51.41%, Est = -0.09, 95% CI [-0.28, 0.10]). The BF for the comparison between this model and a null model that excluded session type (whether the test trial was expected or unexpected) was 0.106[7], therefore favoring the null hypothesis. Our hypothesis-driven model for the amount of time bonobos spent looking at the object (Figure 4.3) similarly revealed neither a significant effect of condition type (ROPE = 3.29%, Est = -0.85, 95% CI [-1.77, 0.07]) nor of session number (ROPE = 22.40%, Est = -0.18, 95% CI [-0.39, 0.03]). The BF for the comparison between this model and the null

---

[7] This means that the hypothesis-driven hypothesis was 0.106 times more likely than the null hypothesis. In other words, the null hypothesis, which excluded session type, was 9.43 times more likely than the hypothesis driven model.

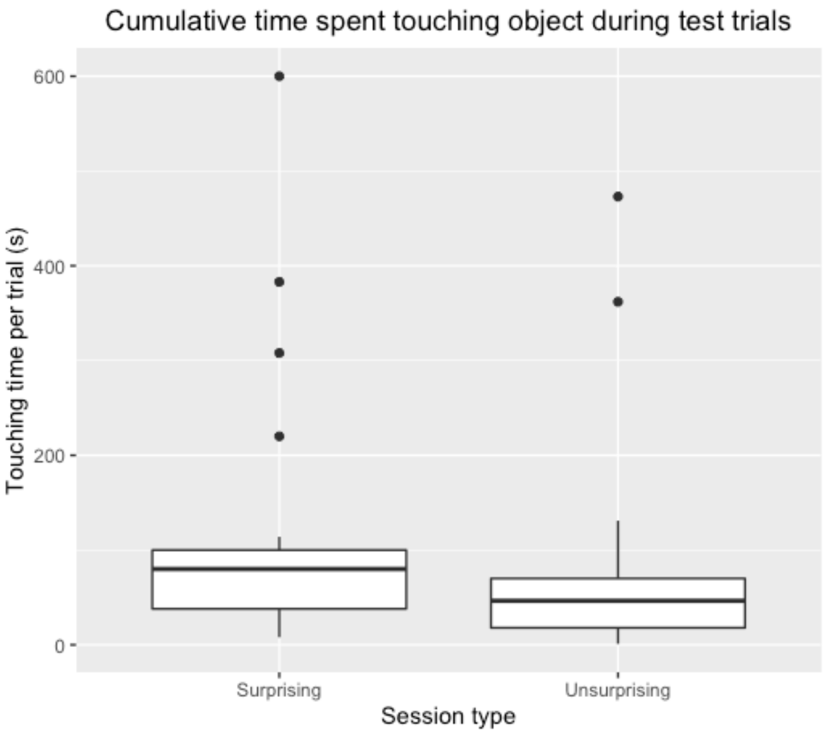

Figure 4.2 Time bonobos spent touching the object during test trials depending on whether they were surprising (i.e. the weight of the object differed from what it weighed during familiarization trial) or unsurprising (i.e. the weight of the object was the same as what it was during the familiarization trial) in Study 4. Boxes demarcate the interquartile ranges, bolded lines demarcate the medians, and points indicate observations that fall outside of the interquartile range.

model that excluded session type was 0.254, revealing that the null hypothesis is more likely.

Lastly, our hypothesis-driven model also found no significant effect of condition type (ROPE = 3.28%, Est = -1.13, 95% CI [-2.54, 0.31]) or of session number (ROPE = 33.17%, Est = -0.14, 95% CI [-0.46, 0.18]) on the number of unique behaviors that bonobos performed on the object (Figure 4.4). The BF for comparing this model to the null model that excluded session type was 0.259, thus favoring the null hypothesis.

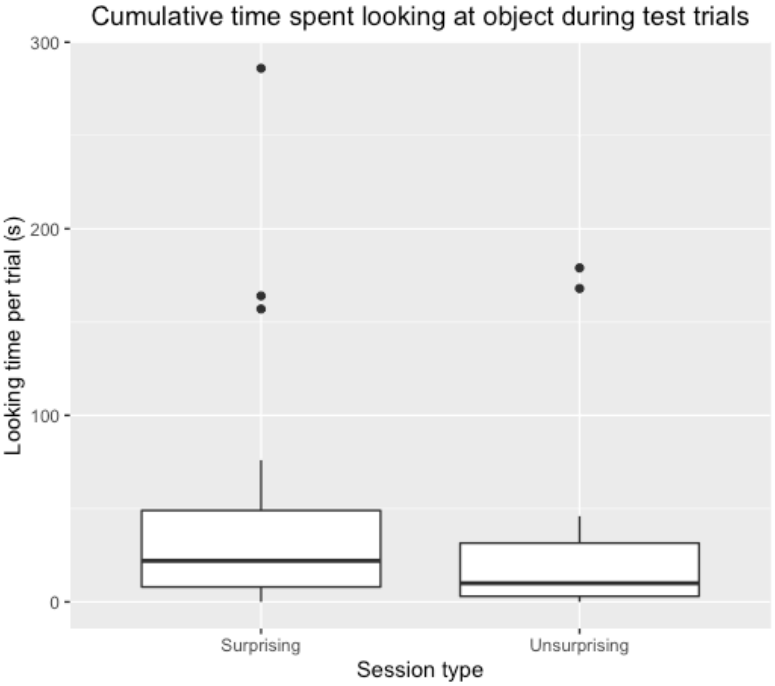

Figure 4.3 How long bonobos looked at the object for, on average, during test trials depending on whether they were surprising or unsurprising in Study 4. Boxes demarcate the interquartile ranges, bolded lines demarcate the medians, and points indicate observations that fall outside of the interquartile range.

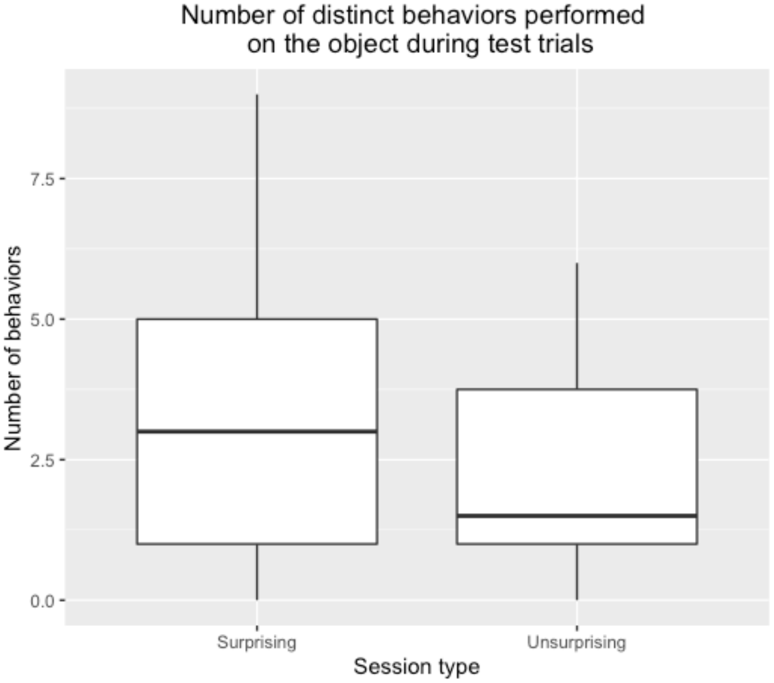

Figure 4.4 Number on unique behaviors bonobos performed on the object for during test trials depending on whether they were surprising or unsurprising in Study 4. Boxes demarcate the interquartile ranges and bolded lines demarcate the medians.

Our exploratory models indicated that neither object type (vase, car, skull, or dinosaur) nor object weight (light or heavy) predicted the amount of time bonobos spent touching or looking at the object or the number of behaviors they performed on it (weight ROPE = 17.25%, Est = 0.16, 95% CI [-0.69, 1.00]). The BFs for the comparison between these models and the null models that removed object type as a factor to assess its impact on touching, looking, and number of behaviors were 0.0005, 0.001, and 0.002, respectively, indicating that the posterior was not more likely given the inclusion of object type in the model. The BFs for the comparison between these models and the null models that removed object weight as a factor to assess its impact on touching, looking, and number of behaviors were 0.045, 0.047, and 0.072, respectively, indicating that including the weight of object in the model did not make the posterior distribution more likely.

4.4.3 Discussion

Study 4 suggests that bonobos did not explore an object significantly more when it surprised them when it did not. We assessed exploration by how much time bonobos spent touching the object, how much time they spent looking at the object, and how many different behaviors they performed on the object. We believed that this would yield insight into a number of different modes of exploration. Amount of time spent touching the object was used as a measure of tactile exploration, and amount of time spent looking at the object was used as a measure of visual exploration. The number of behaviors performed was meant to capture information-seeking, as different behaviors could yield different insights about the object. Our hypothesis-driven analyses revealed that bonobos did not touch or look at the object significantly longer in surprising test trials, nor did they perform significantly more behaviors. In addition, our exploratory analyses revealed that no particular type or weight of object impacted the apes' exploration, and that object type and weight did not account for differences in exploration. The finding that apes did not preferentially explore the unexpected object is consistent with the hypothesis that humans are unique in their explanation-seeking behavior, and that this could serve to scaffold causal reasoning. It is plausibly the case that, as this study indicates, bonobos do not in fact prefer to explore objects that surprise them. However, there are a number of considerations to keep in mind before soundly reaching this conclusion.

While none of these analyses met the threshold of significance for rejecting the null hypothesis, it is worth noting that the trend was in the direction of greater exploration in surprising trials and that ROPE values were quite low. Specifically, the duration of time spent touching the object was 5.38% above the threshold to reject the null hypothesis, and looking at the object and the number of behaviors performed on the object were both less than 0.8% above it. Though it was not a significant effect, it appears that the bonobos tended to touch the object for longer, look at the object for longer, and perform a greater diversity of behaviors on the object when it differed in weight from the familiarization object than when it did not. There are a number of reasons why we may not have reached significance in this study, even if bonobos do in fact explore an object more when it surprises them.

For one, we had a very limited sample size here. Only seven apes participated, and of those only five successfully completed all eight sessions. We also could not compensate for our small sample size with many repeated trials because surprise naturally wanes over repeated exposures. This leads to the second possibility as to why we may not have detected a significant effect; the surprise may have been less effective across repeated sessions. While we did not see an overall trend of session number, there were only a maximum of eight sessions and a maximum of four of these were surprising. Therefore, there may have been an effect of session that we did not have the power to detect. It could be the case that the effect of the surprise became less impactful over subsequent sessions or that the bonobos did not form distinct expectations about the weight of the second object after their first surprising trial when they realize that the objects can be different weights even when they look the same.

Another potential explanation for why we did not find significant evidence of surprise-induced exploration in this population could be due to the fact that they have participated in countless studies throughout their lives, particularly for some individuals in the group. They may therefore 1) suspend some of their expectations in the context of an experiment because they have historically been presented with unusual stimuli (take for example Kanzi, who has been exposed to an immeasurable plethora of human artifacts), or 2) may have interacted with the objects because they recognize that it is part of the task in which they are participating. Bonobos were aware likely aware that they were participating in a study both because they are familiar with the experimenter in the context of research, and because they were separated from their group mates in order to participate. In this sense, touching the object may not be a true measure of interest. One piece of evidence in support of this interpretation is that on multiple occasions the bonobos attempted to pass the object to the experimenter during the test trial. If they were touching the object out of genuine interest, they likely would not be trying to give it away.

For these reasons, future research should be conducted with a much larger sample size and with fewer repetitions per individual. In addition to testing more individuals and a broader range of species, more naturalistic stimuli could be used that would make this study more suitable for a wider range of environments. For example, as in (Sirianni et al., 2018), rocks could be altered such that some are hollow and therefore are much lighter than an ape would expect.

Interacting with rocks would likely be more ecologically valid than colorful man-made objects and might therefore elicit more natural expectations.

Another potential explanation for why we did not find evidence for surprise-induced exploration in bonobos could be rooted in the choice of violation used in this study. While there is evidence that apes (including both human and non-humans) have expectations of the weight of objects, it is possible that violating a different set of expectations might yield a stronger effect. It could be the case that the bonobos do not hold strong expectations about the weight of man-made objects. In this study, we attempted to collect data on the bonobos' hand movements as they lifted the object in order to have a quantitative measure of their expectation of the test objects' weights. The idea was that the bonobos' hands would move upward and with greater acceleration if they expected the object to be heavy and it was actually light, and their hands would move downward if they expected it to be light and it was actually heavy. Unfortunately, the quality of our data did not allow for meaningful analysis. Future work could seek to address this in a more controlled setting, and the objects could be attached to sensors that measure the contact dynamics, as in past research with humans (Buckingham et al., 2009; Flanagan et al., 2008; Gordon et al., 1993; Johansson & Westling, 1988).

Gaining clarity on whether bonobos had specific expectations of the objects' weights would also help to validate that we were in fact examining an effect of surprise here and not just novelty. If the bonobos found the weight-altered objects to be novel, but not surprising per se, this could explain the trend towards longer engagement that did not meet significance. Demery (2013) found that parrots who had been familiarized to an object explored subsequent test objects more when they differed in an external property, such as shape or color, but they explored most the test object differed in an internal property, its center of gravity. They suggest that these external changes made the test objects novel compared to the familiarization object, but the internal change actually evoked surprise because it violated a prior expectation about the object when the parrots engaged with it. Future control studies should be conducted with the bonobos, where apes are presented with objects at test that differ in color or size to see how their exploration compares to when the object at test differs in weight.

Even if bonobos do have strong expectations of an object's weight, this type of violation of expectation may not evoke the motivation to seek an explanation. In future studies, a wider range of surprises should be used to assess whether certain types of violations are more likely than others to illicit surprise-induced exploration. In particular, this violation may have seemed merely improbable, and not altogether impossible, especially since the bonobos saw that there were two objects all along. While it was not possible to implement in the current study because bonobos could easily destroy the familiarization objects, an object seeming to suddenly change an internal property might seem more impossible, and therefore elicit a stronger motivation to seek an explanation, than two objects that appear to be identical having different weights. In addition, this study was designed to be run in parallel with both bonobos and human infants, but we were not able to conduct this comparison due to the suspension of in-person testing because of the COVID-19 pandemic. Once in-laboratory testing resumes, this study should be conducted with human infants to investigate if violations of object weight do indeed incite explanation-seeking in humans.[8]

While we believe that bonobos are an appropriate species to begin to explore explanation-seeking in apes, we see no reason that such studies should be limited to bonobos. Testing a broader range of ape species would create potential for a much larger sample and might

---

[8] Note that we attempted to validate this measure in a study with undergraduates, but the task demands were such that exploration durations could not be meaningful interpreted.

even yield insights into species differences. For one, in the wild chimpanzees are more habitual tool-users than are bonobos (Koops et al., 2015). This could mean that chimpanzees might be more sensitive to the properties of objects, as they must assess them as tools and plan their actions on them. This could also relate to a greater capacity for causal understanding in chimpanzees relative to bonobos (Herrmann et al., 2010), as they engage in action-planning to achieve their goals. Finally, bonobos as a whole tend to be somewhat more neophobic than are other species of apes (Forss et al., 2019; Herrmann et al., 2010), which may make them more hesitant to engage with an unexpected object.

Ultimately, Study 4 does not provide evidence of surprise-induced exploration in bonobos. Of course, this by no means provides conclusive evidence that bonobos do not seek explanations for events that surprise them. This study presents a first step towards addressing the question of whether humans are unique in this regard, and we urge deeper exploration into a much wider range of expectations to be violated and species to be tested. The paradigm presented here can provide a useful template for future studies, and the general design of the task could be used with different sets of expectations in a wider range of environments and species. This study lays a foundation for what we hope will be a fruitful field of research.

Chapter 5 Conclusion

Research with infants and children has shown that young learners engage in exploration when their expectations are violated, seeking explanations for puzzling events (eg: Bonawitz et al., 2012; Chandler & Lalonde, 1994; Perez & Feigenson, 2022; Stahl & Feigenson, 2015; van Schijndel et al., 2015). This form of targeted exploration can serve to help reveal causal mechanisms that might not otherwise be available and can therefore help a learner develop their mental representations of how entities interact in the world. This in turn could help scaffold causal reasoning, the ability to think abstractly about causality, mentally represent opaque causal structures, and predict how objects will interact, even in novel circumstances. Such abilities are essential to humans, as they plan complex action sequences, such as making tools. Primates and other non-human species also have expectations about the world, and research on prediction error has revealed that they, like human infants and children, learn more when these expectations are violated (Pearce & Hall, 1980). Non-human species do not seem to develop the same degree of abstract and flexible reasoning about unseen causal structures, however (Seed et al., 2011). The main question we seek to gain traction on in this thesis is whether humans might be unique in their propensity to seek explanations for surprising events. Throughout this thesis, we present methodologies that were designed to be run in tandem with both human infants and primates to allow for direct comparison. We first aim to sharpen our understanding of the underlying similarities and differences in the expectations that primates hold about how objects act in the world. In Chapter 2, we assessed a commonly-used measure of expectations, looking time, and conclude that this measure may not be as flexible and robust as it is often regarded to be and may not always be the most suitable tool for examining individuals' expectations. Both 9-month-old human infants in an online setting and rhesus macaques in a live, free-ranging setting failed to look longer to events that past research has indicated should be surprising to both species. In Chapter 3, we presented a new paradigm for assessing understanding of object properties, and we provide evidence that bonobos can use their expectations about how entities interact to reason about the location of objects, even when they are out of sight. Finally, in Chapter 4, we examined surprise-induced exploration in bonobos and provided preliminary evidence that they do not prioritize surprising objects for exploration. In other words, bonobos do not appear to seek explanations for unexpected events. In this chapter, we briefly summarize the framework and findings from each of these chapters, identify open questions that remain, and suggest directions for future work to begin to gain traction into these questions.

5.1 Summary of Chapters 1-4

5.1.1 Validating looking time as a measure of expectancy violation in human infants and primates

The studies in Chapter 2 aimed to evaluate looking time as a measure of violation of expectations in two distinct contexts: with human infants tested online and with rhesus macaques tested with live stimuli in a naturalistic context. While looking time has long been considered the gold standard for evaluating non-verbal individuals' expectations (see Stahl & Kibbe, 2022 for review), we provided evidence that this measure may not be as robust and flexible as it is often regarded to be. Using stimuli modeled directly after those from previous studies (Baillargeon et al., 1992; Spelke et al., 1992; Stahl & Feigenson, 2015), we failed to evoke the violation of expectations (VOE) effect in both populations. Human infants and rhesus macaques were no more likely to look at an object that appeared to roll through a wall than stop when it hit a wall (Studies 1 and 2, respectively), nor did they look longer at an object that appeared to stop in midair as opposed to remaining in contact with a supporting surface (Study 1). We speculated on the possible causes of these failures and ultimately concluded that looking time to violations of core knowledge of objects may not be as sensitive when infants are tested on a 2-dimensional display and when monkeys are tested with inedible objects as compared to food items. While these findings point to potential limitations of the use of looking time to evaluate infants' and primates' expectations, we do not take these findings as evidence of a lack of expectations of core knowledge of objects in infants and primates.

5.1.2 Bonobos' use of core knowledge about the location of hidden food

The study in Chapter 3 aimed to present a new paradigm for assessing expectations of core knowledge in bonobos. While apes are presumed to have expectations of objects that are rooted in core knowledge, based on evidence of such expectations in species that are more distantly related to humans (eg in dogs: Kundey et al., 2010; chicks: Regolin & Vallortigara, 1995; rhesus macaques: Santos & Hauser, 2002), very little research has directly tested such expectations. This study went beyond just asking whether apes recognize a violation of core knowledge of objects when it occurs, and instead asked whether bonobos can *use* their expectations of object properties to reason about the location of food items they cannot see. Consistent with past research with human children (Butler et al., 2002; Hood et al., 2000) and primates (Cacchione et al., 2009; Hauser, 2001; Hauser et al., 2001; Hood et al., 1999; Tomonaga et al., 2007), bonobos were successful in utilizing knowledge about the properties of solidity and spatiotemporal continuity to make predictions about the location of objects, but they struggled to reason as flexibly about object support. Study 3 provided evidence that apes can reason about unseen entities using principles of core knowledge of objects and can use this understanding to plan their actions.

5.1.3 Surprise-induced exploration in bonobos

The work in Chapter 4 began to address the question of whether apes, like humans, engage in explanation-seeking upon experiencing an unexpected event. Human infants and children have been shown to preferentially explore objects that surprise them (Bonawitz et al., 2012; Chandler & Lalonde, 1994; Charlesworth, 1964; Perez & Feigenson, 2022; Povinelli & Dunphy-Lelii, 2001; Sim & Xu, 2017; Stahl & Feigenson, 2015; van Schijndel et al., 2015), to tailor their exploration to the type of violation they have witnessed (Stahl & Feigenson, 2015), and to conduct informative manipulations that isolate variables to maximize the information gained through exploration (Legare, 2012; van Schijndel et al., 2015). We presented apes with objects that appeared to be identical to one another but had been manipulated to differ significantly in an internal property of the object, its weight. Past research indicates that apes are sensitive to the weight of objects and form expectations about object weight (Sirianni et al., 2018), and research with parrots demonstrated that they preferred to explore objects that had an internal property altered (Demery, 2013). It was therefore reasonable to believe that apes would form expectations about the weight of the objects and would be surprised when the weight of the object was altered. The results from Study 4 did not provide evidence of surprise-induced exploration and explanation-seeking in bonobos, as they did not spend significantly longer touching or looking at the surprising objects as compared to the unsurprising ones, nor did they perform significantly more behaviors on these objects. We discuss potential alternative explanations for why apes may not have preferentially explored the objects that changed weight and emphasize the importance of future research that aims to extend this study to a larger sample size, a wider range of species, and a wider range of violations.

5.2 Open question: Do apes engage in explanation seeking to a wider range of violations?

In Study 4, we presented evidence that bonobos did not seem to seek an explanation for an object having a different weight than they would expect. Even with this finding, we are far from concluding that apes do not engage in surprise-induced exploration. We strongly encourage future research that aims to replicate Study 4 with a larger sample and with a wider range of species. In particular, we think that it would be especially informative to replicate this study with chimpanzees, as they are more avid tool-users in the wild (Koops et al., 2015) and may therefore engage in more nuanced action planning, where sensitivity to object properties would be more relevant. There is also past research that demonstrates that chimpanzees have expectations about the weight of objects that they use as tools and adjust their behaviors in accordance with the objects' weights (Sirianni et al., 2018). Therefore, we think the current paradigm would be well-suited for replication with chimpanzees. Such a study could also yield insight into the evolutionary pressures that might shape the propensity to seek explanations for surprising events. If chimpanzees, but not bonobos, engage in such surprise-induced exploration, this could be an indication that the complexity of action planning, particularly in the context of tool use, is a relevant factor in the development of this learning mechanism. Similar comparisons could be conducted with other closely-related taxa that do an do not use tools, such as tufted capuchins as compared to white-faced capuchins (Ottoni & Izar, 2008).

In addition, a wider range of expectations should be used in subsequent work. It could be the case that bonobos do not have strong expectations that two identical man-made objects should weigh the same, or that a violation of such an expectation is not one that triggers a search for an explanation. Future research could capitalize on expectations that we know that bonobos hold. For example, one could show bonobos an object that appears to violate the principles of spatiotemporal continuity or solidity, which Study 3 demonstrates are expectations that bonobos hold, and then make the object available for exploration. We believe that this line of research is only in its infancy and is ripe for further elaboration and development.

5.3 What is the role of the subjective experience of surprise?

In this thesis we review evidence that human infants and children preferentially learn about and explore objects that surprise them. All of these studies involve showing participants something that surprises them and then evaluating the viewer's response. This therefore conflates the cognitive process of expectation being violated with the subjective experience of surprise. In the research presented in this thesis, we fail to find evidence for such explanation seeking in one of humans' two closest living relatives. At this point, it known that primates experience the cognitive processes of surprise, but it is unknown if they experience the subjective experience of surprise (Kret et al., 2020). One pathway to determining the potential underpinnings of surprise-induced exploration that might differentiate humans from our primate relatives would be to assess the role of the subjective experience of surprise in prioritizing an object for subsequent investigation.

We suggest that a starting point to begin to address this question is to assess whether infants can use someone else's surprise as a cue that there is something to learn. Wu et al. (2019)

provide evidence that 12-17-month-olds can use someone else's reaction to make predictions about the outcomes of unseen events, in what they term "vicarious prediction error". Thus, it is plausible that infants can use someone else's surprise as a cue that an instance is worth prioritizing for subsequent learning. If infants succeeded on such a task, this would provide evidence that the subjective experience of surprise is not necessary to trigger surprise-induced learning and exploration, and it is therefore plausible to expect apes to engage in surprise-induced learning even if they do not in fact experience the *feeling* of surprise, upon witnessing a violation of expectations.

5.4 Open question: What is the mechanism that links explanation-seeking to causal reasoning?

Throughout this thesis, we present the argument that exploring an object that violates one's expectations can help a learner to reveal previously opaque causal mechanisms, and that individuals can then update their mental representation of the causal dynamics of the event. We argue that through this process, surprise-induced exploration can serve to scaffold causal reasoning, as learners can come to think more abstractly about how entities relate. At this stage, we are lacking concrete evidence of this link. We certainly believe that such a process is plausible, but future research should examine this question directly. Specifically, are infants more likely to grasp the underlying causal mechanism of an unexpected event after exploring? Studies that have examined surprise-induced learning have demonstrated that the new information learners glean from the surprising need not be causal, as infants more easily learned the sound that the surprising object made (Stahl & Feigenson, 2015), and preschoolers more easily learned the word for a surprising object (Stahl & Feigenson, 2017). This makes clear that instances of surprise are prioritized for learning, but it is not yet clear that the information to be gained is specifically causal. Future studies should directly test whether exploring the target of an unexpected event makes children more likely to reveal the underlying causal mechanism of the event and whether their understanding of future events seems to have been informed by this updated mental representation.

5.5 Conclusion

In this thesis, we presented evidence that humans and primates do not seem to differ in their underlying expectations of the physical world, but that humans may be unique in their propensity to actively seek explanations when those expectations are violated. In this way, humans may reveal causal mechanisms that would not otherwise be apparent. They can then begin to update their mental model of the world and begin to think more abstractly about how inanimate entities relate to one another. This propensity to explore the things that surprise us may have arisen during the course of evolution, as humans engaged in greater degrees of tool use and performed behaviors that require greater intricacy of action planning and understanding of cause and effect. The work presented in this thesis provides a first step in examining surprise-induced exploration in non-humans, and it begins to lay the foundation for an area that is ripe for further exploration.

References

Baillargeon, R. (1987). Object permanence in 3½- and 4½-month-old infants. *Developmental Psychology*, *23*(5), 655–664. https://doi.org/10.1037/0012-1649.23.5.655

Baillargeon, R. (1995). Physical Reasoning in Infancy. *The Cognitive Neurosciences*, 181–204.

Baillargeon, R., Needham, A., & Devos, J. (1992). The development of young infants' intuitions about support. *Early Development and Parenting*, *1*(2), 69–78. https://doi.org/10.1002/edp.2430010203

Baldwin, D. A., & Markman, E. M. (1993). Infants' ability to draw inferences about nonobvious object properties: Evidence from exploratory play. *Child Development*, *64*(3), 711–728. https://doi.org/10.2307/1131213

Barr, R. (2013). Memory Constraints on Infant Learning From Picture Books, Television, and Touchscreens. *Child Development Perspectives*, *7*(4), 205–210. https://doi.org/10.1111/cdep.12041

Barr, R., & Hayne, H. (1999). Developmental Changes in Imitation from Television during Infancy. *Child Development*, *70*(5), 1067–1081.

Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or Surprise? *Frontiers in Psychology*, *4*. https://www.frontiersin.org/article/10.3389/fpsyg.2013.00907

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting Linear Mixed-Effects Models using lme4* (arXiv:1406.5823). arXiv. https://doi.org/10.48550/arXiv.1406.5823

Berthier, N. E., DeBlois, S., Poirier, C. R., Novak, M. A., & Clifton, R. K. (2000). Where's the ball? Two- and three-year-olds reason about unseen events. *Developmental Psychology*, *36*(3), 394–401. https://doi.org/10.1037/0012-1649.36.3.394

Bochynska, A., & Dillon, M. R. (2021). Bringing Home Baby Euclid: Testing Infants' Basic Shape Discrimination Online. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/article/10.3389/fpsyg.2021.734592

Bonawitz, E. B., Ferranti, D., Saxe, R., Gopnik, A., Meltzoff, A. N., Woodward, J., & Schulz, L. E. (2010). Just do it? Investigating the gap between prediction and action in toddlers' causal inferences. *Cognition*, *115*(1), 104–117. https://doi.org/10.1016/j.cognition.2009.12.001

Bonawitz, E. B., van Schijndel, T. J. P., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, *64*(4), 215–234. https://doi.org/10.1016/j.cogpsych.2011.12.002

Buchsbaum, D., Bridgers, S., Skolnick Weisberg, D., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2202–2212. https://doi.org/10.1098/rstb.2012.0122

Buckingham, G., Cant, J. S., & Goodale, M. A. (2009). Living in A Material World: How Visual Cues to Material Properties Affect the Way That We Lift Objects and Perceive Their Weight. *Journal of Neurophysiology*, *102*(6), 3111–3118. https://doi.org/10.1152/jn.00515.2009

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In *The developmental psychology of time* (pp. 209–254).

Bürkner, P.-C. (2021). Bayesian Item Response Modeling in R with brms and Stan. *Journal of Statistical Software*, *100*, 1–54. https://doi.org/10.18637/jss.v100.i05

Busch, J. T. A., & Legare, C. H. (2019). Using data to solve problems: Children reason flexibly in response to different kinds of evidence. *Journal of Experimental Child Psychology*, *183*, 172–188. https://doi.org/10.1016/j.jecp.2019.01.007

Butler, S. C., Berthier, N. E., & Clifton, R. K. (2002). Two-year-olds' search strategies and visual tracking in a hidden displacement task. *Developmental Psychology*, *38*(4), 581–590. https://doi.org/10.1037/0012-1649.38.4.581

Cacchione, T., Call, J., & Zingg, R. (2009). Gravity and solidity in four great ape species (Gorilla gorilla, Pongo pygmaeus, Pan troglodytes, Pan paniscus): Vertical and horizontal variations of the table task. *Journal of Comparative Psychology*, *123*(2), 168–180. https://doi.org/10.1037/a0013580

Cacchione, T., & Krist, H. (2004). Recognizing Impossible Object Relations: Intuitions About Support in Chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *118*(2), 140–148. https://doi.org/10.1037/0735-7036.118.2.140

Call, J. (2010). Trapping the minds of apes: Causal knowledge and inferential reasoning about object-object interactions. In E. V. Lonsdorf, S. R. Ross, T. Matsuzawa, & J. Goodall (Eds.), *The Mind of the Chimpanzee: Ecological and Experimental Perspectives* (pp. 75–86). University of Chicago Press.

Call, J., & Tomasello, M. (2005). Reasoning and Thinking in Nonhuman Primates. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 601–632).

Chandler, M. J., & Lalonde, C. E. (1994). Surprising, magical and miraculous turns of events: Children's reactions to violations of their early theories of mind and matter. *British*

Journal of Developmental Psychology, *12*(1), 83–95. https://doi.org/10.1111/j.2044-835X.1994.tb00620.x

Chappell, J., Demery, Z. P., Arriola-Rios, V., & Sloman, A. (2012). How to build an information gathering and processing system: Lessons from naturally and artificially intelligent systems. *Behavioural Processes*, *89*(2), 179–186. https://doi.org/10.1016/j.beproc.2011.10.001

Charlesworth, W. R. (1964). Instigation and Maintenance of Curiosity Behavior as a Function of Surprise versus Novel and Familiar Stimuli. *Child Development*, *35*(4), 1169–1186. https://doi.org/10.2307/1126863

Chiandetti, C., & Vallortigara, G. (2010). Experience and geometry: Controlled-rearing studies with chicks. *Animal Cognition*, *13*(3), 463–470. https://doi.org/10.1007/s10071-009-0297-x

Chuey, A., Asaba, M., Bridgers, S., Carrillo, B., Dietz, G., Garcia, T., Leonard, J. A., Liu, S., Merrick, M., Radwan, S., Stegall, J., Velez, N., Woo, B., Wu, Y., Zhou, X. J., Frank, M. C., & Gweon, H. (2021). Moderated Online Data-Collection for Developmental Research: Methods and Replications. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/article/10.3389/fpsyg.2021.734398

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*(3), 341–349. https://doi.org/10.1016/j.cognition.2011.03.003

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*(7), 294–300. https://doi.org/10.1016/j.tics.2006.05.004

Croker, S., & Buchanan, H. (2011). Scientific reasoning in a real-world context: The effect of prior belief and outcome on children's hypothesis-testing strategies. *British Journal of Developmental Psychology*, *29*(3), 409–424. https://doi.org/10.1348/026151010X496906

Csink, V., Mareschal, D., & Gliga, T. (2021). Does surprise enhance infant memory? Assessing the impact of the encoding context on subsequent object recognition. *Infancy*, *26*(2), 303–318. https://doi.org/10.1111/infa.12383

Darwin, C. (1872). *The expression of the emotions in man and animals*. John Murray.

Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core Knowledge of Geometry in an Amazonian Indigene Group. *Science*, *311*(5759), 381–384. https://doi.org/10.1126/science.1121739

Demery, Z. P. (2013). *Comparative sensory & cognitive adaptations for exploratory learning in parrots & humans* [D_ph, University of Birmingham]. https://etheses.bham.ac.uk/id/eprint/4526/

Diener, M. L., Pierroutsakos, S. L., Troseth, G. L., & Roberts, A. (2008). Video Versus Reality: Infants' Attention and Affective Responses to Video and Live Presentations. *Media Psychology*, *11*(3), 418–441. https://doi.org/10.1080/15213260802103003

Ekman, P., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions* (pp. xiv, 496). Oxford University Press.

Everett, D. L. (2005). Cultural Constraints on Grammar and Cognition in Pirahã Another Look at the Design Features of Human Language. *Current Anthropology*, *46*(4), 621–646. https://doi.org/10.1086/431525

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, *16*(1), 88–92. https://doi.org/10.3758/PBR.16.1.88

Feigenson, L., Carey, S., & Hauser, M. (2002). The Representations Underlying Infants' Choice of More: Object Files Versus Analog Magnitudes. *Psychological Science*, *13*(2), 150–156. https://doi.org/10.1111/1467-9280.00427

Flanagan, J. R., Bittner, J. P., & Johansson, R. S. (2008). Experience Can Change Distinct Size-Weight Priors Engaged in Lifting Objects and Judging their Weights. *Current Biology*, *18*(22), 1742–1747. https://doi.org/10.1016/j.cub.2008.09.042

Forss, S. I. F., Motes-Rodrigo, A., & Tennie, C. (2019). Animal Behavior: Ape Curiosity on Camera. *Current Biology*, *29*(7), R255–R257. https://doi.org/10.1016/j.cub.2019.02.055

Girndt, A., Meier, T., & Call, J. (2008). Task constraints mask great apes' ability to solve the trap-table task. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(1), 54–62. https://doi.org/10.1037/0097-7403.34.1.54

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. *Psychological Review*, *111*(1), 3–32. https://doi.org/10.1037/0033-295X.111.1.3

Gopnik, A., & Tenenbaum, J. B. (2007). Bayesian networks, Bayesian learning and cognitive development. *Developmental Science*, *10*(3), 281–287. https://doi.org/10.1111/j.1467-7687.2007.00584.x

Gordon, A. M., Westling, G., Cole, K. J., & Johansson, R. S. (1993). Memory representations underlying motor commands used during manipulation of common and novel objects. *Journal of Neurophysiology*, *69*(6), 1789–1796. https://doi.org/10.1152/jn.1993.69.6.1789

Gordon, P. (2004). Numerical Cognition Without Words: Evidence from Amazonia. *Science*, *306*(5695), 496–499. https://doi.org/10.1126/science.1094492

Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, *94*, 149–165. https://doi.org/10.1016/j.jml.2016.11.001

Gweon, H., & Schulz, L. (2011). 16-Month-Olds Rationally Infer Causes of Failed Actions. *Science*, *332*(6037), 1524–1524. https://doi.org/10.1126/science.1204493

Hanus, D., & Call, J. (2008). Chimpanzees infer the location of a reward on the basis of the effect of its weight. *Current Biology*, *18*(9), R370–R372. https://doi.org/10.1016/j.cub.2008.02.039

Hare, B., Wobber, V., & Wrangham, R. (2012). The self-domestication hypothesis: Evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour*, *83*(3), 573–585. https://doi.org/10.1016/j.anbehav.2011.12.007

Hauser, M. D. (2001). Searching for food in the wild: A nonhuman rimate's expectations about invisible displacement. *Developmental Science*, *4*(1), 84–93. https://doi.org/10.1111/1467-7687.00152

Hauser, M. D., & Carey, S. (2003). Spontaneous representations of small numbers of objects by rhesus macaques: Examinations of content and format. *Cognitive Psychology*, *47*(4), 367–401. https://doi.org/10.1016/S0010-0285(03)00050-1

Hauser, M. D., Williams, T., Kralik, J. D., & Moskovitz, D. (2001). What guides a search for food that has disappeared? Experiments on cotton-top tamarins (Saguinus oedipus). *Journal of Comparative Psychology*, *115*(2), 140–151. https://doi.org/10.1037/0735-7036.115.2.140

Herrmann, E., Hare, B., Call, J., & Tomasello, M. (2010). Differences in the Cognitive Skills of Bonobos and Chimpanzees. *PLOS ONE*, *5*(8), e12438. https://doi.org/10.1371/journal.pone.0012438

Hood, B., Carey, S., & Prasada, S. (2000). Predicting the Outcomes of Physical Events: Two-Year-Olds Fail to Reveal Knowledge of Solidity and Support. *Child Development*, *71*(6), 1540–1554. https://doi.org/10.1111/1467-8624.00247

Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a non-human primate? *Developmental Science*, *2*(1), 35–41. https://doi.org/10.1111/1467-7687.00051

Hopper, L. M., Gulli, R. A., Howard, L. H., Kano, F., Krupenye, C., Ryan, A. M., & Paukner, A. (2021). The application of noninvasive, restraint-free eye-tracking methods for use with nonhuman primates. *Behavior Research Methods*, *53*(3), 1003–1030. https://doi.org/10.3758/s13428-020-01465-6

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

Horstmann, G. (2006). The time course of intended and unintended allocation of attention. *Psychological Research*, *70*(1), 13–25. https://doi.org/10.1007/s00426-004-0184-7

Horstmann, G., Becker, S., & Ernst, D. (2016). Perceptual salience captures the eyes on a surprise trial. *Attention, Perception, & Psychophysics*, *78*(7), 1889–1900. https://doi.org/10.3758/s13414-016-1102-y

Horstmann, G., & Herwig, A. (2015). Surprise attracts the eyes and binds the gaze. *Psychonomic Bulletin & Review*, *22*(3), 743–749. https://doi.org/10.3758/s13423-014-0723-1

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. https://doi.org/10.1016/j.visres.2008.09.007

Johansson, R. S., & Westling, G. (1988). Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Experimental Brain Research*, *71*(1), 59–71. https://doi.org/10.1007/BF00247522

Kay, M. (2022). *tidybayes: Tidy Data and Geoms for Bayesian Models*. Zenodo. https://doi.org/10.5281/zenodo.5823492

Keen, R. (2003). Representation of Objects and Events: Why Do Infants Look So Smart and Toddlers Look So Dumb? *CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE*, *12*(3), 30.

Koops, K., Furuichi, T., & Hashimoto, C. (2015). Chimpanzees and bonobos differ in intrinsic motivation for tool use. *Scientific Reports*, *5*(1), 11356. https://doi.org/10.1038/srep11356

Kret, M. E., Prochazkova, E., Sterck, E. H. M., & Clay, Z. (2020). Emotional expressions in human and non-human great apes. *Neuroscience & Biobehavioral Reviews*, *115*, 378–395. https://doi.org/10.1016/j.neubiorev.2020.01.027

Kundey, S. M. A., De Los Reyes, A., Taglang, C., Baruch, A., & German, R. (2010). Domesticated dogs' (Canis familiaris) use of the solidity principle. *Animal Cognition*, *13*(3), 497–505. https://doi.org/10.1007/s10071-009-0300-6

Kushnir, T., & Gopnik, A. (2005). Young Children Infer Causal Strength From Probabilities and Interventions. *Psychological Science*, *16*(9), 678–683. https://doi.org/10.1111/j.1467-9280.2005.01595.x

Lea, S. E. G., Slater, A. M., & Ryan, C. M. E. (1996). Perception of object unity in chicks: A comparison with the human infant. *Infant Behavior and Development*, *19*(4), 501–504. https://doi.org/10.1016/S0163-6383(96)90010-7

Legare, C. H. (2012). Exploring Explanation: Explaining Inconsistent Evidence Informs Exploratory, Hypothesis-Testing Behavior in Young Children. *Child Development*, *83*(1), 173–185. https://doi.org/10.1111/j.1467-8624.2011.01691.x

Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency With Prior Knowledge Triggers Children's Causal Explanatory Reasoning. *Child Development*, *81*(3), 929–944. https://doi.org/10.1111/j.1467-8624.2010.01443.x

Legare, C. H., Schult, C. A., Impola, M., & Souza, A. L. (2016). Young children revise explanations in response to new evidence. *Cognitive Development*, *39*, 45–56. https://doi.org/10.1016/j.cogdev.2016.03.003

Legare, CristineH., & Gelman, SusanA. (2014). Examining Explanatory Biases in Young Children's Biological Reasoning. *Journal of Cognition & Development*, *15*(2), 287–303. https://doi.org/10.1080/15248372.2012.749480

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9

Lewis, L. S., & Krupenye, C. (2022). Eye-tracking as a window into primate social cognition. *American Journal of Primatology*, e23393. https://doi.org/10.1002/ajp.23393

Limongelli, L., Boysen, S. T., & Visalberghi, E. (1995). Comprehension of cause€ffect relations in a tool-using task by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *109*(1), 18–26. https://doi.org/10.1037/0735-7036.109.1.18

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Martin-Ordas, G., & Call, J. (2009). Assessing generalization within and between trap tasks in the great apes. *International Journal of Comparative Psychology*, *22*(1), 19.

Martin-Ordas, G., Call, J., & Colmenares, F. (2008). Tubes, tables and traps: Great apes solve two functionally equivalent trap tasks but show no evidence of transfer across tasks. *Animal Cognition*, *11*(3), 423–430. https://doi.org/10.1007/s10071-007-0132-1

McClelland, J. L. (2002). Prediction-error driven learning: The engine of change in cognitive development. *Proceedings 2nd International Conference on Development and Learning. ICDL 2002*, 43-. https://doi.org/10.1109/DEVLRN.2002.1011732

McCormack, T., Bramley, N., Frosch, C., Patrick, F., & Lagnado, D. (2016). Children's use of interventions to learn causal structure. *Journal of Experimental Child Psychology*, *141*, 1–22. https://doi.org/10.1016/j.jecp.2015.06.017

Mulcahy, N. J., & Call, J. (2006). How great apes perform on a modified trap-tube task. *Animal Cognition*, *9*(3), 193–199. https://doi.org/10.1007/s10071-006-0019-6

Munakata, Y., Santos, L. R., Spelke, E. S., Hauser, M. D., & O'Reilly, R. C. (2001). Visual Representation in the Wild: How Rhesus Monkeys Parse Objects. *Journal of Cognitive Neuroscience*, *13*(1), 44–58. https://doi.org/10.1162/089892901564162

Munnich, E., Ranney, M. A., & Song, M. (2007). *Surprise, Surprise: The Role of Surprising Numerical Feedback in Belief Change*. 7.

Murai, C., Tanaka, M., & Sakagami, M. (2011). Physical intuitions about support relations in monkeys (Macaca fuscata) and apes (Pan troglodytes). *Journal of Comparative Psychology*, *125*(2), 216–226. https://doi.org/10.1037/a0022099

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition*, *47*(2), 121–148. https://doi.org/10.1016/0010-0277(93)90002-D

O'Connell, S., & Dunbar, R. I. M. (2005). The perception of causality in chimpanzees (Pan spp.). *Animal Cognition*, *8*(1), 60–66. https://doi.org/10.1007/s10071-004-0231-1

Ottoni, E. B., & Izar, P. (2008). Capuchin monkey tool use: Overview and implications. *Evolutionary Anthropology: Issues, News, and Reviews*, *17*(4), 171–178. https://doi.org/10.1002/evan.20185

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. https://doi.org/10.1037/0033-295X.87.6.532

Penn, D. C., & Povinelli, D. J. (2007). Causal Cognition in Human and Nonhuman Animals: A Comparative, Critical Review. *Annual Review of Psychology*, *58*(1), 97–118. https://doi.org/10.1146/annurev.psych.58.110405.085555

Perez, J., & Feigenson, L. (2022). Violations of expectation trigger infants to search for explanations. *Cognition*, *218*, 104942. https://doi.org/10.1016/j.cognition.2021.104942

Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*(3), 302–321. https://doi.org/10.1016/j.cognition.2010.11.015

Pieraut-Le Bonniec, G. (1985). From visual-motor anticipation to conceptualization: Reaction to solid and hollow objects and knowledge of the function of containment. *Infant Behavior and Development*, *8*(4), 413–424. https://doi.org/10.1016/0163-6383(85)90005-0

Povinelli, D. J., & Dunphy-Lelii, S. (2001). Do chimpanzees seek explanations? Preliminary comparative investigations. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *55*(2), 185–193. https://doi.org/10.1037/h0087365

Premack, D., & Premack, A. J. (1994). Levels of causal understanding in chimpanzees and children. *Cognition*, *50*(1), 347–362. https://doi.org/10.1016/0010-0277(94)90035-3

Regolin, L., & Vallortigara, G. (1995). Perception of partly occluded objects by young chicks. *Perception & Psychophysics*, *57*(7), 971–976. https://doi.org/10.3758/BF03205456

Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The Cognitive-Evolutionary Model of Surprise: A Review of the Evidence. *Topics in Cognitive Science*, *11*(1), 50–74. https://doi.org/10.1111/tops.12292

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. *Classical Conditioning: Current Research and Theory*.

Retell, J. D., Venini, D., & Becker, S. I. (2015). Oculomotor Capture by New and Unannounced Color Singletons during Visual Search. *Attention, Perception, & Psychophysics*, *77*(5), 1529–1543. https://doi.org/10.3758/s13414-015-0888-3

Santos, L., & Hood, B. (2009). Object representation as a central issue in cognitive science. *The Origins of Object Knowledge*. https://doi.org/10.1093/acprof:oso/9780199216895.003.0001

Santos, L. R. (2004). ‘Core Knowledges’: A dissociation between spatiotemporal knowledge and contact-mechanics in a non-human primate? *Developmental Science*, *7*(2), 167–174. https://doi.org/10.1111/j.1467-7687.2004.00335.x

Santos, L. R., & Hauser, M. D. (2002). A non-human primate’s understanding of solidity: Dissociations between seeing and acting. *Developmental Science*, *5*(2), F1–F7. https://doi.org/10.1111/1467-7687.t01-1-00216

Schmitt, K. L. (1997). *Two- to three-year-olds’ understanding of the correspondence between television and reality* [Ph.D., University of Massachusetts Amherst]. http://www.proquest.com/docview/304353299/abstract/6DDFACFEBB144393PQ/1

Schulz, L. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences*, *16*(7), 382–389. https://doi.org/10.1016/j.tics.2012.06.004

Schulz, L. (2015). Infants explore the unexpected. *Science*, *348*(6230), 42–43. https://doi.org/10.1126/science.aab0582

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*(4), 1045–1050. https://doi.org/10.1037/0012-1649.43.4.1045

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology*, *40*(2), 162–176. https://doi.org/10.1037/0012-1649.40.2.162

Schulz, L. E., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science*, *10*(3), 322–332. https://doi.org/10.1111/j.1467-7687.2007.00587.x

Schulz, L. E., Standing, H. R., & Bonawitz, E. B. (2008). Word, thought, and deed: The role of object categories in children's inductive inferences and exploratory play. *Developmental Psychology*, *44*(5), 1266–1276. https://doi.org/10.1037/0012-1649.44.5.1266

Seed, A., Hanus, D., & Call, J. (2011). Causal Knowledge in Corvids, Primates, and Children. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool Use and Causal Cognition* (pp. 89–110). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199571154.003.0005

Seed, A. M., Call, J., Emery, N. J., & Clayton, N. S. (2009). Chimpanzees solve the trap problem when the confound of tool-use is removed. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(1), 23–34. https://doi.org/10.1037/a0012925

Sim, Z. L., & Xu, F. (2017). Infants preferentially approach and explore the unexpected. *British Journal of Developmental Psychology*, *35*(4), 596–608. https://doi.org/10.1111/bjdp.12198

Sirianni, G., Wittig, R. M., Gratton, P., Mundry, R., Schüler, A., & Boesch, C. (2018). Do chimpanzees anticipate an object's weight? A field experiment on the kinematics of hammer-lifting movements in the nut-cracking Taï chimpanzees. *Animal Cognition*, *21*(1), 109–118. https://doi.org/10.1007/s10071-017-1144-0

Smith-Flores, A. S., Perez, J., Zhang, M. H., & Feigenson, L. (2022). Online measures of looking and learning in infancy. *Infancy*, *27*(1), 4–24. https://doi.org/10.1111/infa.12435

Sobel, D. M., & Munro, S. E. (2009). Domain generality and specificity in children's causal inference about ambiguous data. *Developmental Psychology*, *45*(2), 511–524. https://doi.org/10.1037/a0014944

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science*, *28*(3), 303–333. https://doi.org/10.1207/s15516709cog2803_1

Southgate, V., & Gomez, J. C. (2006). Searching beneath the shelf in macaque monkeys: Evidence for a gravity bias or a foraging bias? *Journal of Comparative Psychology*, *120*(3), 314–321. https://doi.org/10.1037/0735-7036.120.3.314

Spelke, E. S. (2000). Core knowledge. *American Psychologist*, *55*(11), 1233–1243. https://doi.org/10.1037/0003-066X.55.11.1233

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605–632. https://doi.org/10.1037/0033-295X.99.4.605

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Spelke, E. S., & Lee, S. A. (2012). Core systems of geometry in animal minds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1603), 2784–2793. https://doi.org/10.1098/rstb.2012.0210

Stahl, A. E. (2015). *The role of surprise in enhancing early learning*. 126.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science (New York, N.Y.)*, *348*(6230), 91–94. https://doi.org/10.1126/science.aaa3799

Stahl, A. E., & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Cognition*, *163*, 1–14. https://doi.org/10.1016/j.cognition.2017.02.008

Stahl, A. E., & Feigenson, L. (2019). Violations of Core Knowledge Shape Early Learning. *Topics in Cognitive Science*, *11*(1), 136–153. https://doi.org/10.1111/tops.12389

Stahl, A. E., & Kibbe, M. M. (2022). Great expectations: The construct validity of the violation-of-expectation method for studying infant cognition. *Infant and Child Development*. https://doi.org/10.1002/icd.2359

Subbotsky, E. (2010). Curiosity and exploratory behaviour towards possible and impossible events in children and adults. *British Journal of Psychology*, *101*(3), 481–501. https://doi.org/10.1348/000712609X470590

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, *10*(7), 309–318. https://doi.org/10.1016/j.tics.2006.05.009

Tomasello, M., Call, J., & Call, P. of P. Y. P. C. J. (1997). *Primate Cognition*. Oxford University Press.

Tomkins, S. (1962). *Affect Imagery Consciousness: Volume I: The Positive Affects*. Springer Publishing Company.

Tomonaga, M., Imura, T., Mizuno, Y., & Tanaka, M. (2007). Gravity bias in young and adult chimpanzees (Pan troglodytes): Tests with a modified opaque-tubes task. *Developmental Science*, *10*(3), 411–421. https://doi.org/10.1111/j.1467-7687.2007.00594.x

Troseth, G. L., & DeLoache, J. S. (1998). The Medium Can Obscure the Message: Young Children's Understanding of Video. *Child Development*, *69*(4), 950–965. https://doi.org/10.1111/j.1467-8624.1998.tb06153.x

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The Automaticity of Visual Statistical Learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Uller, C., Carey, S., Hauser, M., & Xu, F. (1997). Is language needed for constructing sortal concepts? A study with nonhuman primates. *Proceedings of the 21st Annual Boston University Conference on Language Development*, *21*.

Uller, C., Hauser, M., & Carey, S. (2001). Spontaneous representation of number in cotton-top tamarins (Saguinus oedipus). *Journal of Comparative Psychology*, *115*(3), 248–257. https://doi.org/10.1037/0735-7036.115.3.248

Vaesen, K. (2012). The cognitive bases of human tool use. *Behavioral and Brain Sciences*, *35*(4), 203–218. https://doi.org/10.1017/S0140525X11001452

Valenza, E., Leo, I., Gava, L., & Simion, F. (2006). Perceptual Completion in Newborn Human Infants. *Child Development*, *77*(6), 1810–1821. https://doi.org/10.1111/j.1467-8624.2006.00975.x

van Schijndel, T. J. P., Visser, I., van Bers, B. M. C. W., & Raijmakers, M. E. J. (2015). Preschoolers perform more informative experiments after observing theory-violating evidence. *Journal of Experimental Child Psychology*, *131*, 104–119. https://doi.org/10.1016/j.jecp.2014.11.008

Visalberghi, E., & Limongelli, L. (1994). Lack of comprehension of cause€ffect relations in tool-using capuchin monkeys (Cebus apella). *Journal of Comparative Psychology*, *108*(1), 15–22. https://doi.org/10.1037/0735-7036.108.1.15

Visalberghi, E., & Tomasello, M. (1998). Primate causal understanding in the physical and psychological domains. *Behavioural Processes*, *42*(2), 189–203. https://doi.org/10.1016/S0376-6357(97)00076-4

Völter, C. J., Lambert, M. L., & Huber, L. (2020). Do Nonhumans Seek Explanations? *Animal Behavior and Cognition*, *7*(3), 445–451. https://doi.org/10.26451/abc.07.03.10.2020

Walden, T., Kim, G., McCoy, C., & Karrass, J. (2007). Do you believe in magic? Infants' social looking during violations of expectations. *Developmental Science*, *10*(5), 654–663. https://doi.org/10.1111/j.1467-7687.2007.00607.x

Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition*, *93*(3), 167–198. https://doi.org/10.1016/j.cognition.2003.09.012

Wu, Y., Merrick, M., & Gweon, H. (2019). *Infants Use Others' Surprise as Vicarious Prediction Error*. PsyArXiv. https://doi.org/10.31234/osf.io/8whuv