THE DEVELOPMENT OF CAUSAL EXPLANATORY REASONING

by

Cristine H. Legare

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in The University of Michigan
2008

Doctoral Committee:

Professor Susan A. Gelman, Co-Chair
Professor Henry M. Wellman, Co-Chair
Professor Peter A. Railton
Research Scientist Evelyn Margaret Evans

Acknowledgements

This research was supported by the McDonnell Collaborative Initiative on Causal Learning. Thanks to the children, parents, and staff at Pooh Corner Preschool and the University of Michigan Children's Center. I am grateful to my friends (and fellow students): Tamar Kushnir, Andrea Kiss, Jenny LaBounty, Medha Tare, Besangie Sellars, Sonja Baumer, and many more. Special thanks to Paul Theron, for your boundless support and enthusiasm for me and my work.

I am grateful to Henry Wellman for very helpful advice, many wonderful conversations, and for guiding me through the process of academic writing, Margaret Evans for her careful reading and challenging questions, and Peter Railton for his great insight. Thank you very much for your support and interest in my work.

To Gail Heyman, thank you for your time, thoughtful advice, and endless encouragement. Thank you also for inspiring me by your success, and making me hopeful about the future.

And most importantly, Susan Gelman, thank you for your truly exceptional mentorship, and for supporting me in everything I wanted to do – from studying AIDS in South Africa to internships with UNICEF. Thank you for your guidance and encouragement, for believing in me and investing in my work. I am very fortunate to have had such a wonderful example of both personal and professional accomplishment.

I am grateful for my supportive, loving family. A special thanks to my siblings, Carolyn, Charles, and Clayton, and my mother, Diane.

Table of Contents

List of Tables

List of Figures

Abstract

Young children actively seek to understand the world around them; they construct causal explanations for how and why things happen. The early-developing capacity for causal explanatory reasoning raises several questions: How do children assemble causal-explanatory systems of knowledge? What motivates children to construct causal explanations? What can the kinds of events that trigger causal explanatory reasoning tell us about the function of children's explanations?

In a series of studies with preschool children, contrastive outcomes were used as an experimental paradigm for studying the kinds of events that provoke children's causal explanations. In Study 1 (N=48, age range 3,2 to 5,6) and Study 2 (N=32, age range 3,0 to 4,11), in order to investigate two competing hypotheses about the function of children's explanations, events that were inconsistent with children's prior knowledge were simultaneously contrasted with events that were consistent with children's prior knowledge. Results suggest that inconsistent outcomes are an especially powerful trigger for children's explanations, and that children provide explanations for inconsistent outcomes that refer to underlying, internal causal properties, overriding perceptual appearances.

Study 3(N=28 children, age range 3,1 to 5,2; N=16 adults) specifically targeted state-change and negative outcomes as additional kinds of explanatory triggers, within a knowledge-rich context (illness). In Study 3, preschool children's causal reasoning about illness was investigated, specifically, their explanations for *preventing* illness versus

curing illness. Results indicate that state-change and negative outcomes provoke children's causal explanations. As predicted, illness prevention provokes explanations less often than illness cure or treatment.

In sum, data provide evidence for the interplay of three distinct, but interrelated biases that guide children's causal explanatory reasoning. The data also provide insight into the function of children's explanations and empirical evidence for the kinds of events that motivate children to construct explanations.

Chapter I

Introduction: The Function of Children's Causal Explanations

From early childhood onward, children actively work to understand the world around them; they seek to explain how and why things happen. The capacity for causal explanatory reasoning raises several questions: In general, how do children assemble causal-explanatory systems of knowledge? Given the sheer number of events children could attend to and attempt to understand, what motivates children to construct specific causal explanations? That is, what are the events, outcomes, confusions, and goals that trigger children's explanations? Finally, what can these causal explanatory triggers tell us about the function of children's explanations?

Children make use of causal-explanatory understanding to explain *consistent events* (when events unfold as anticipated based on prior knowledge) but also to recognize and attempt to explain *inconsistent events* (when something unusual or discordant with prior knowledge happens). Accordingly, children's explanations may serve at least two distinct functions. One possibility is that explanation serves as a mechanism for confirming children's prior knowledge. Children are early in the process of developing explanatory knowledge and are faced with the considerable task of navigating an infinite number of outcomes and events that could potentially warrant explanation. Therefore, consolidating and confirming their explanations for events consistent with prior knowledge and experience may be especially attractive and beneficial.

Additionally, children readily make use of covariation information, statistical regularities, and causal relationships in order to understand causal outcomes, frequently from very limited available input (Gopnik, Sobel, Schulz, & Glymour, 2001; Schulz & Gopnik, 2004; Kushnir & Gopnik, 2007). If children have a cognitive model of the world based on a framework of anticipatory causal regularities, they would be well-equipped to rapidly form expectations contingent upon prior beliefs or knowledge. Given a predisposition to prognosticate causal regularities, children may anticipate that *outcomes will continue to occur as expected* and find consistent outcomes especially worthy of explanation. Constructing explanations for events that are consistent with children's prior knowledge and experience may indeed be an important function of children's own explanations. For example, explaining consistent outcomes may provide children with an important opportunity to deepen their understanding of causal phenomena by allowing them to generate causal mechanisms.

Another intriguing possibility is that explanation is motivated by discovery. That is, young children might especially value, seek, and provide explanations for events that are inconsistent with their current expectations. According to this possibility, because children readily form expectations for causal regularities based on prior knowledge (even when sparse), children may be highly motivated to attend to irregular or discordant information. Information that is inconsistent with how they expect things to happen could be especially informative and noteworthy because it indicates that their prior knowledge about a causal relationship or outcome was incomplete or inaccurate. Therefore, children may be vigilantly attentive to, and importantly, more likely to attempt to explain disconfirmatory outcomes. If this were the case, engaging in explanation would allow children the opportunity to accommodate and reconcile inconsistent information in the context of prior beliefs, and forming explanations for inconsistent outcomes may provide children with the opportunity to generate new hypotheses regarding events that seem to disconfirm their prior knowledge.

Before addressing the question of how to empirically test these two competing hypotheses about the function of children's explanations, I will present an overview of the developmental literature on causal explanatory reasoning, followed by a discussion of the role of prior knowledge in shaping causal explanation, and an overview of the kinds of events likely to provoke or trigger causal explanation. Finally, methodological approaches to investigating potential explanatory biases experimentally will be presented.

The development of causal reasoning: The role of explanation

Although the development of causal reasoning has been an important topic in developmental psychology since Piaget (1929), children's causal reasoning has received renewed attention in more recent years (Gopnik & Schulz, 2007), especially from those characterizing children's knowledge in terms of naïve theories (Carey, 1985; Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1994; Keil, 1989; Wellman & Gelman, 1998).

A substantial amount of developmental research has demonstrated that young children understand many general causal principles (Bullock, Gelman, & Baillargeon, 1982; Kushnir & Gopnik, 2005; Schulz & Gopnik, 2004; Shultz, 1982) and possess remarkably rich causal knowledge (Wellman & Gelman, 1998). However, less is known about how causal reasoning develops and the role explanation plays in this process.

One of the primary reasons that the Piagetian account of children's causal knowledge has been replaced in the last three decades is a shift in methods for assessing children's understanding. Piagetian methods depended heavily on analyzing children's explanations, which have been criticized for underestimating children's knowledge (Bullock, Gelman, & Baillargeon, 1982). As Bullock et al. (1982) noted, "Children's explanations for events did not seem to reflect the same level of causal reasoning as did their judgments or predictions…..The results are, of course, not a surprise to anyone working with preschool-age children. Children are more likely to demonstrate their reasoning in actions and simple choices than explanations" (p. 246). As a result, most contemporary research with young children has focused instead on judgment tasks that ask for predictions.

It seems reasonable to assume that causal predictions could be less demanding and emerge earlier than causal explanations. Causal predictions can be based on detecting causal regularities whereas causal explanations typically require conceptualizing an outcome relative to a more general framework of interpretation. Furthermore, predictions can be manifest in simple yes/no or behavioral judgments whereas causal explanations typically require more extended verbal expression and reasoning. However, although traditional Piagetian investigations failed to portray children's abilities accurately, I propose that it is not explanation itself that is problematic (see also Wellman & Liu, 2007). Intriguingly, research in the psychological and biological domains has found that preschool children provide pertinent explanations in advance of accurate predictions (Amsterlaw & Wellman, 2006; Bartsch & Wellman, 1989, Legare, Wellman, & Gelman, under review), thus showing that young children's explanations can be surprisingly revealing.

A central function of causal reasoning is to provide explanations for phenomena in the world. Causal explanations play a central role in both everyday reasoning (Gopnik, 2000; Hickling & Wellman, 2001; Hilton, 1988; Keil, 2000; 2006; Sloman, 2005) and scientific theories (Hempel, 1965; Pitt, 1988; Salmon, 1984, 1989; Strevens, 2006; Trout, 2002, 2007; Woodward, 2003). Additionally, prominent theories of conceptual development (Carey, 1985; Keil, 1995, 2003) and category learning (Murphy, 2002; Murphy & Allopenna, 1994) assign a central role to causal-explanatory understanding, claiming that explanation is central to the nature and development of naïve theories (Wellman, 1990) or that it is characteristic of all concepts (Murphy & Medin, 1985). Relatedly, in the philosophy of science, successful theories are those that are consistently explanatorily successful (Railton, 1989).

Indeed, the explanatory component of developing knowledge structures may be especially crucial. Children's causal explanations both demonstrate their understandings of the world and, like their questions (Chouinard, 2007), may constitute a mechanism for advancing causal learning and the acquisition of knowledge (Amsterlaw & Wellman, 2006; Bartsch & Wellman, 1989; Callanan & Oakes, 1992; Gopnik & Meltzoff, 1997; Siegler, 1995).

Emerging developmental research has begun to focus on explanation (Frazier, Gelman, & Wellman, under review; Keil, 2006; Keil & Wilson, 2000; Legare, Wellman, & Gelman, under review). To the extent that explicit "why" questions and "because" answers can be used as prototypical indices of explanatory reasoning, developmental research indicates that both explanations and requests for explanation are widespread even in very young children. Research examining preschoolers' everyday conversations with their caregivers has demonstrated the frequency of causal-explanatory utterances by young children (Callanan & Oakes, 1992; Hickling & Wellman, 2001). Causal explanations increase in frequency with age but are common even at 2-3 years of age (Wellman, Hickling, & Schult, 1997). Furthermore, causal explanations most typically serve an epistemic function; that is, they provide an interpretation for a current or past event, and do not serve an exclusively social-regulatory function (Hickling & Wellman, 2001). Given the proliferation of explanatory activity young children engage in: What triggers or motivates children to generate causal explanations?

Causal explanation is a goal-directed human activity. It depends on what is relevant or important to the person constructing an explanation. A desire to understand may underlie the motivation to construct an explanation (Gopnik, 1996). According to Gopnik (1996, 2000), the phenomenology or experience of explanation is an essential component of the task of explanation. One possibility is that a drive to explain evolved because generally speaking, it aids in learning and contributes to an increasingly accurate understanding of the causal structure of the world around us. A strong interest in constructing explanations may be especially beneficial for learning in childhood.

Prior knowledge and causal explanation

Because explanations are contingent to some extent upon prior knowledge (Medin, Coley, & Storms, 2003; Sloman, 1994), engaging in explanation serves as a mechanism through which prior knowledge or beliefs are brought to bear on a relevant inference or judgment, serving as a basis for interpreting an inconsistent or novel outcome. Additionally, prior knowledge constrains causal inference and explanation by reducing the range of possible causal mechanisms considered to be explanatorily informative or relevant.

Although the kind of information or evidence that children and scientists have access to is importantly different (Brewer, Chinn, & Samarapungavan, 2000; Cummins, 2000), the thesis that prior knowledge plays a central role in shaping children's explanations is consistent with models of causal explanation in the philosophy of science literature (Glymour, 2000). For example, the Deductive-Nomological (DN) model of explanation indicates that, "the essence of scientific explanation can be described as *nomic expectability* — that is expectability on the basis of lawful connections" (Salmon, 1989, p. 57).

Although the *DN* model is meant to capture explanation via deduction from deterministic natural laws, unlike scientific explanation in physics and chemistry, most causal explanations in everyday life cannot be deduced from laws, at least not laws that meet the standard criteria for lawfulness (Cummins, 2000).   In most cases of non-scientific causal explanation, both statistical regularities and causal mechanisms play a central role (Ahn & Kalish, 2000). Both philosophical (Hume, 1777) and psychological theories of causal reasoning taking the regularity view emphasize that causal strength is a function of a covariation index and statistical regularities (Cheng, 1997; Cheng & Novick, 1992), whereas proponents of process or mechanism approaches emphasize the central role of the transmission of causal influence in reasoning about causation (Ahn & Kalish, 2000; Ahn, Kalish, Medin, & Gelman, 1995; Salmon, 1984).

However, normative models of causal explanation in science can be interpreted as supporting logical arguments in support of either of the proposed functions of causal explanation in children, namely, explanation as confirmation *or* explanation as discovery (Cummins, 2000). That is, although the literature of scientific explanation clearly supports the centrality of expectations based on prior knowledge, explanation could be motivated by (a) increasing understanding of causal mechanisms underlying information consistent with prior knowledge or (b) reconciling and accommodating information inconsistent with prior knowledge.

Research in cognitive psychology also provides evidence that causal explanation is intimately tied to prior knowledge (Keil, 2006; Keil & Wilson, 2000; Lombrozo, 2006). Although traditional models of causal inference have emphasized covariation or other measures of statistical evidence (Shanks, 1995), recent research with adults has shown that the interpretation and impact of such evidence depends on prior beliefs (Ahn, Marsh, & Luhmann, 2007). Additionally, there are data demonstrating that explanations of covariation between a candidate cause and effect can determine whether covariation is taken as evidence for causation (Koslowski, 1996) and that prior causal knowledge influences learning and inference (Koslowski & Thompson, 2002; Tenenbaum, Griffiths, & Kemp, 2006).

Explanatory triggers

The aim of this dissertation was to investigate causal explanatory biases or triggers in preschool children by using their explanations as the primary dependent measure. Although the idea that inconsistent, problematic, or surprising outcomes play an important role in causal reasoning appears across multiple literatures--philosophy of science (Hempel, 1965), social psychology (Hilton, 1995), educational research (Chi, Bassok, Lewis, Reimann, Glaser, 1989), and infancy research (Baillargeon, 2002)--there is remarkably little empirical research on what motivates causal explanations in children and how this can inform the developmental trajectory of causal explanation. Investigating this question requires appropriate controls, in addition to clear alternatives.

Because young children have so much to learn, they have much to explain. How do children navigate the task of causal learning and what motivates children to construct causal explanations? I propose that explanatory biases play an important role in guiding children's causal explanations. Although the sheer number of events or outcomes young children could be interested in explaining is considerable, the overarching objective of this dissertation was to investigate three interrelated, but conceptually distinct kinds of outcomes that may potentially trigger causal explanation in children: outcomes inconsistent with prior knowledge, state-change, and negative outcomes. If children anticipate regularity or consistency with prior beliefs, they may therefore find *outcomes inconsistent with prior knowledge* especially worthy of explanation. For example, children may expect an object to continue to function or a person to continue to behave in a manner consistent with prior experience. Alternative functioning or anomalous behavior would therefore be inconsistent outcomes.

Additionally, if children anticipate that current perceptible states will not change, they may be especially compelled to explain *state-change* outcomes. For example, based on prior knowledge, children may expect individuals to stay healthy or objects to remain in motion. State-changes, such as recovery from illness, may therefore constitute a specific kind of inconsistency with prior experience. Although state-change often corresponds to inconsistency with prior knowledge, it isn't always the case. However, a perceptible change in the physical state of an outcome may be highly compelling and prompt explanation even in the absence of any prior experience.

Finally, children may anticipate or assume that outcomes will be positive or favorable, in which case *negative outcomes* would trigger children's explanations. That is, children may be more compelled to explain a disappointment or loss than an achievement. Like state-change, negative outcomes also often correspond to inconsistent outcomes. For example, if prior experience indicates that outcomes will generally be favorable, a negative experience (such as poor performance on a task) would be inconsistent.

For these reasons, although state-change and negative outcomes frequently correspond to inconsistency with prior knowledge, they can be conceptually distinct. They may also be heuristically useful and compelling in the absence of prior experience or background knowledge. For example, attending to and explaining physical transformations and threatening events could help children acquire important and instructive new information. Because it is possible that all three kinds of outcomes may provoke causal explanation in children, and because these kinds of outcomes are often confounded both in the world and in experimental manipulations, a primary motivation behind the design of Studies 1-3 was to experimentally differentiate the kinds of events that children find noteworthy and therefore feel compelled to explain.

Contrastive outcomes, counterfactuals, and causation

In order to investigate the kinds of events children are most compelled to explain, I turn now to a discussion of when children's explanations involve invoking a contrast case, as opposed to an exclusive focus on the event-to-be-explained. Counterfactual thinking entails mentally comparing the observed case with alternative cases, and this process may especially direct attention to and provoke interest in inconsistent outcomes. When observing an event, children build a representation of the event that they use to interpret, explain, and predict its outcomes. Causal judgments often involve a contrast between a perceived sequence and a counterfactual case (Hilton & Slugoski, 1986), and developmental research indicates that preschool children use counterfactual thinking in causal reasoning (Harris, German, & Mills, 1996). That is, the identification of a new outcome often involves noticing contrastive outcomes and identifying the conditions that are causally responsible for differences between the outcomes. Contrary to the Humean account of causal learning, Mackie (1974) argues that our beliefs about causality are not based exclusively on repeated observations but also on an interpretation of what is observed and what might have been observed instead. Therefore, what we describe as a cause is an antecedent condition that is determined to play a causal role in relation to a specific event or set of circumstances. Had the antecedent condition not occurred, neither would the outcome. For example, when pondering an explanation for why the car broke down, one may determine that it would have continued to run if there had been sufficient gas in the tank.

Interestingly, although research using contrastive outcomes with children is limited (but see Harris, et al., 1996), research on infant cognition relies heavily on the use of contrastive outcome tasks as a way to prime infants' expectations, thereby suggesting, implicitly, that contrastive outcomes may be an important way to assess young children's understanding. Infant cognition research also provides support for the hypothesis that inconsistent or problematic outcomes are compelling from a very early age (Wang, Baillargeon, & Brueckner, 2004). Violation-of-expectation (VOE) tasks have been widely used to assess infants' understanding of physical (Baillargeon, 2002) and psychological (Onishi, Baillargeon, & Leslie, 2007) phenomena, based on the assumption that infants: (a) have expectations, (b) are surprised when these expectations are violated, and (c) index surprise by showing greater attention as determined by increased looking time. In a typical VOE experiment, infants watch two test events, one consistent with the expectation examined in the experiment (expected event) and one inconsistent (unexpected event). In order to introduce potentially unfamiliar test stimuli or establish specific expectations, prior to the test trials, infants usually view habituation or familiarization trials. With appropriate controls, evidence that infants look reliably longer at the unexpected than at the expected event is taken to indicate that they possess the expectation under investigation, detect the violation in the expected event, and are surprised by this violation (Wang et al., 2004). Surprise is defined as synonymous with a state of heightened attention or arousal caused by an expectation violation.

However, it is difficult to tell whether these responses from infants are truly surprise, or even expectation-violation. And with infants, one cannot tell for sure if they are actively exploring or genuinely seeking more information. For example, there is no good evidence that longer looking time corresponds to other measures of emotional state (Haith & Benson, 1997; Haith, 1998; Wang et al., 2004). Although claims about explanatory phenomenology and expectation violation have been amply made in the infancy literature, there is still a big gap between the behaviors that can be measured (such as looking time) and the phenomenon of theoretical interest (for present purposes, whether an event is in need of explanation). Therefore, examining explanatory reasoning with older children may provide us with more concrete insights for several reasons. Children's causal explanations do more than demonstrate their understandings of the world through verbal articulation: they may also constitute a mechanism for advancing causal learning and the acquisition of knowledge (Lombrozo, 2006). Moreover, because children are actively engaged in developing causal knowledge structures, constructing explanations may engage their emerging curiosity and understanding. Additionally, prototypical indices of explanatory reasoning such as "why" questions and "because" answers indicate that both explanations and requests for explanation are widespread even in 2-year-olds.

Present studies

The objective of this dissertation was to explore explanatory triggers in preschool children by investigating their explanations. In the following studies, contrastive outcomes were used as an experimental paradigm for studying the kinds of events that provoke children's causal explanations. Because the kinds of events that motivate causal explanatory reasoning have implications for the function of children's explanations, the objective of Studies 1 and 2 was to investigate whether children reason differently about consistent events versus events in which something inconsistent or unexpected happens. If explanation is largely confirmatory, children should be motivated to construct explanations for outcomes that are consistent with prior knowledge. If, on the other hand, explanation is a mechanism for discovery, children should be motivated to construct explanations for events that are inconsistent with their prior knowledge. I predict that outcomes inconsistent with prior knowledge are a powerful explanatory trigger for children's causal reasoning.

Study 3 expands upon Studies 1 and 2 in three ways by (a) extending the findings from Studies 1 and 2 to a new domain, (b) specifically targeting state-change and negative outcomes as additional kinds of explanatory triggers, and (c) investigating explanatory triggers in a knowledge-rich context.

Chapter II

Study 1: Investigating Outcomes Inconsistent with Prior Knowledge as a Trigger for Causal Explanatory Reasoning in Young Children

"Observation is always selection. It needs a chosen object, a definite task, an interest, a point of view, a problem."

-Popper, 1963

A fundamental task for all humans is explaining why things happen. Research on conceptual development indicates that even children as young as 3 years of age can use causal knowledge to make predictions (Shultz, 1982), engage in efficacious interventions (Kushnir & Gopnik, 2007; Schulz & Gopnik, 2004), and provide explanations for phenomena in the world (Wellman, Hickling, & Schult, 1997). Not only do young children frequently seek explanations by asking questions (Callanan & Oakes, 1992; Choinard, 2007; Hickling & Wellman, 2001), they also construct their own explanations (Frazier, Gelman, & Wellman, under review).

However, other research (e.g., concerning children's metacognition) has shown that young children are surprisingly poor at assessing their own understanding; this ability develops dramatically across development. Indeed, both adults and children overestimate the detail and depth of their explanatory knowledge (Mills & Keil, 2004;

Wilson & Keil, 1998). Taken together, these two sets of findings produce something of a paradox in the literature on children's causal reasoning. On the one hand, children are active explanation-seekers and readily seek out and provide causal explanations. On the other hand, they seem to be poor at assessing their own causal knowledge and often think they understand things when they do not. What then motivates children to ask questions and generate explanations, if they are often concluding that they understand something when they do not? More specifically, what kinds of events provoke causal explanatory reasoning in children? What kinds of events do children feel most compelled to explain?

Children make use of causal-explanatory understanding to explain *consistent events* (when events unfold as anticipated based on prior knowledge) but also to recognize and attempt to explain *inconsistent events* (when something unusual or not in accord with prior knowledge happens). Accordingly, children's explanations may serve at least two distinct functions. One possibility is that explanation is largely confirmatory. That is, because children are early in the process of developing explanatory knowledge, consolidating and confirming their explanations for events is especially attractive (and useful). On this possibility, because children are making use of covariation information, statistical regularities, and causal relationships in order to understand causal outcomes (Gopnik, Sobel, Schulz, & Glymour, 2001; Schulz & Gopnik, 2004; Kushnir & Gopnik, 2007), they may anticipate that based on prior beliefs or knowledge *outcomes will continue to occur as expected* and find such outcomes especially easy and worthy of explanation. Given this kind of expectation, children may be especially motivated to attend to and provide explanations for information that is consistent with how they expect things to happen. Constructing explanations for events that are consistent with children's prior knowledge and experience may indeed be an important function of children's own explanations. For example, explaining consistent outcomes may provide children with an important opportunity to deepen their understanding of causal phenomena by allowing children to generate causal mechanisms.

Another intriguing possibility is that explanation is motivated by discovery. That is, young children might especially value, seek, and provide explanations for events that are inconsistent with their current expectations. Forming explanations for inconsistent outcomes may provide children with the opportunity to generate new hypotheses regarding events that seem to disconfirm their prior knowledge. According to this possibility, based on their expectations for causal regularity, children may be vigilantly attentive to information that is inconsistent with how they expect things to happen. If this were the case, children may not only attend to but also attempt to explain these disconfirmatory outcomes. Therefore, engaging in explanation would allow children the opportunity to accommodate inconsistent information in the context of prior beliefs by providing a basis for interpreting inconsistent events or outcomes.

Additionally, finding specific solutions to meaningful problems may be more compelling than explaining how things generally work (Ahn & Kalish, 2000). For example, understanding and explaining how a can opener normally works entails a different cognitive process than explaining how or why one is broken or how to fix it.

When something goes wrong, or simply does not go as predicted, there is an element of the unexpected, a problem to be solved.

A third possibility is that neither inconsistent nor consistent events are primary explanatory triggers for children, but instead, other kinds of outcomes are responsible for provoking children's explanations. For example, outcomes with an end-state perceptibly different from the initial-state, or state-change outcomes may be especially compelling. Outcomes with a negative valence may also be noteworthy from an explanatory perspective.

In scientific theorizing, disconfirming events play a special role in provoking explanations. Do children engage in an analogous process? I propose that (a) events inconsistent with prior knowledge are especially powerful triggers for explanatory reasoning and (b) events consistent with prior knowledge are less likely to motivate children to construct explanations. Although it is not a novel idea to suggest that inconsistent events spark curiosity (Wang et al., 2004), or even that for adults that they spark attempts to explain them (Simon, 2001), it is worth pointing out that there is little empirical research on what motivates young children to construct explanations and the function of children's own explanations.

How might these alternative possibilities be tested? Imagine that a child sees two equivalent events, one in accord with prior knowledge and the other not. If explanation is largely confirmatory, children should simply explain what they already have an explanation for. If explanation is instead responsive to discordant or anomalous information, children should explain the event that falls outside their prior knowledge or expectations. This is the scenario I use as an experimental paradigm to examine these competing hypotheses about the function of children's explanations. Moreover, I examined the nature of children's explanations, specifically whether children provided explanations primarily in terms of surface features and past histories, or whether they offered explanations focused on less-obvious properties.

To address these issues experimentally, I designed a task with a set of novel "light boxes" – electronic devices which glowed bright when activated. The activation and deactivation of the boxes were experimenter-controlled, but appeared to be caused by objects placed on the surface of each box (materials were modeled after those used in Gopnik & Sobel, 2000; Gopnik et al., 2001). These materials were used to teach children about different categories of objects, where within each category items were both perceptually identical and shared common causal properties. In this first study, objects were labeled according to their causal properties: "starters" were objects which activated the light box when placed on top of it, "stoppers" were objects which deactivated the light box when placed on top of it, and "do-nothings" were objects which could neither activate nor deactivate the light box. After training, children were presented with scenarios in which a new object which looked like one type (for example, it looked like a "starter") actually behaved like another type (behaved like a "do-nothing"). This was paired with an object which looked and behaved like those previously seen (looked like a "do-nothing" and behaved like a "do-nothing"). Upon viewing such paired outcomes, children were asked a non-specific explanatory question ambiguously referring to either (visible) outcomes: "Why did that happen?"

An important feature of the design of this study was experimentally differentiating the kinds of events that children find noteworthy and therefore feel compelled to explain. Although investigating outcomes inconsistent with prior knowledge was of primary interest in this study, state-change and negative outcomes are hypothesized to function as additional potential explanatory triggers. Therefore two focal conditions were designed in order to control for these factors by holding them constant. In the generative condition, both light boxes turned on in one of the test trials and remained off in the other test trial. Therefore, in the test trial in which both boxes turned on, both outcomes involved state-change and neither outcome involved a negative outcome (defined here as failing to turn the boxes on). In the parallel condition in which both boxes stayed off, neither outcome involved a state-change, and both outcomes involved a negative outcome (both failed to turn the boxes on). In the inhibitory condition, this pattern was reversed, see Table 1.

An additional feature of this study was to include a confirmation trial prior to the test trials. The advantage of the confirmation trial was that it allowed for isolating change-of-state outcomes from outcomes inconsistent with prior knowledge. For example, in the confirmation trials, because both outcomes work as expected based on prior experience, "starters" and "stoppers" changed the state of the box and "do-nothings" did not change the state of the box.

The events which led to greater interest and attention, and most importantly, increased explanation, were examined. If the role of explanations for children is confirmatory, children should be interested in and provide explanations for the consistent event in this pair. If "anomaly" plays a special role they could be specifically interested in and provide hypotheses and explanations for the inconsistent event. Figure 4 illustrates the pattern of data that would be anticipated based on these distinct hypotheses about the function of children's explanations.

Given that children were provided with information about the objects' functions and labels, what kinds of explanations might be anticipated? Would children refer to surface appearances, make inferences about underlying causal properties and past histories, or refer to category membership in their explanations for outcomes?

Even children as young as 3 years of age can categorize objects in terms of novel, non-obvious properties (Jaswal & Markman, 2007; Graham, Kilbreath, & Welder, 2004; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987), apply names to objects with the same functional properties (Kemler-Nelson, 1995), and categorize and name objects based on novel causal properties (Gopnik & Sobel, 2000), overriding perceptual appearances. Whether and how preschool children might make use of this information in their causal explanations (as opposed to predictions or judgments) is an open question. Therefore, an additional objective for Study 1 was to investigate whether children's explanations typically include information about function and causal properties. If information about an object's underlying causal properties or function is central to how children categorize and reason about objects, information about function and underlying causal properties should be found in children's causal explanations.

Method

Participants

Sixteen 3-year-olds (*M* age 3,6; range 3,2 to 3,11), sixteen 4-year-olds (*M* age 4,6; range 4, 0 to 4,11), and sixteen 5-year-olds (*M* age 5,3; range 5,0 to 5,6) were recruited from a Midwestern university town, and were primarily White. Approximately equal numbers of boys and girls participated in the study.

Materials

Four "light boxes" (5 inches by 5 inches by 5 inches) were made of wood with thin translucent laminate wooden tops. The boxes were identical except that two were painted red and two yellow. Each box was connected to an electrical outlet and a switch box. Each pair of boxes (red or yellow) was attached to the same switch box and was operated surreptitiously and out-of-sight by a confederate. If the switchbox was turned on, the box would light up and stay on until it was switched off. Alternately, if the switchbox was turned off, the box would turn off and stay off until it was switched on. Each box was turned off or on as soon as an object made contact with it, and would stay off or on until the object was removed. This yielded the strong impression that some objects turned the boxes on, some turned them off, and some had no effect on the boxes. The switchbox was hidden from the children's view, and none of the participants mentioned the confederate or indicated any suspicion that the confederate influenced the functioning of the boxes.

There were four perceptually distinct sets of objects, each consisting of five identical wooden blocks (20 blocks total). These were also always presented in paired sets. There was no relationship between the observed causal properties of the blocks (lighting or not lighting the boxes) and their perceptual features (Figure 1). Additionally, the placement of the objects on the boxes was counterbalanced in order to prevent potential effects of the perceptive shape of the objects on the probability of attracting attention and eliciting explanations.

Procedure

Children were tested individually in a 15-minute session. The child was seated across from the experimenter and confederate at a table on which a pair of the wooden boxes was placed. Each child participated in two conditions: a generative condition and an inhibitory condition. In the generative condition, the boxes (either yellow or red, counterbalanced across children) started with the light off, and the two distinctive sorts of objects were labeled either as “starters” or “do-nothings” (Figure 2). In the inhibitory condition, the boxes started with the light on, and the objects were labeled either as “stoppers” or “do-nothings” (Figure 3). After the objects’ functions and labels were presented individually, both objects were contrasted simultaneously in the confirmation trial. For example, in the generative condition, after demonstrating how a starter and a do-nothing work for one pair of objects, the experimenter placed a different starter and do-nothing simultaneously on two unlit boxes. In this confirmation trial, both objects worked in a manner consistent with prior knowledge. That is, upon contact with one box, the starter turned the light on and the do-nothing had no effect (the light remained off) on the other box. To provide a baseline comparison for their subsequent test-trial explanations, children were asked to explain this initial paired event. That is, with both objects on their boxes, one box on and one off, children were asked: “Why did that happen?”

In the following test trials, two contrasting objects were again simultaneously placed on the boxes. One object worked in a manner consistent with prior knowledge and the other object functioned in a manner that was inconsistent. Note that unlike the confirmation trial, in the test trials, the outcomes in both cases were causally identical.

Both boxes lit up in one test trial and both boxes did not light up in the other test trial. Thus, to us, and perhaps to the child, some objects caused or failed to cause their normative outcome. For example, in a generative cause case (Figure 2), a starter failed to start its target device, or in an inhibitory case, a stopper did not turn off its device (Figure 3). After viewing the paired outcomes of each trial, children were again asked the non-specific explanatory question referring to either visible outcome, "Why did that happen?" Children were prompted several times in a non-directed manner after the additional causal question to provide additional explanatory information. For example, children were asked, "can you tell me more" or "do you have any other ideas" after each trial. The child was encouraged to provide as many explanations as possible. The trial was not terminated after the child's first explanation but only after the child indicated that they had no other information to share.

Because state-change and negative outcomes are hypothesized to function as additional potential explanatory triggers, an important feature of the design of Study 1 was to control for these factors by holding them constant. For example, in the test trial in the generative condition in which both boxes turned on, both outcomes involved state-change and neither outcome involved a negative outcome (defined here as failing to turn the boxes on). In the parallel condition in which both boxes stayed off, neither outcome involved a state-change, and both outcomes involved a negative outcome (both failed to turn the boxes on). In the inhibitory condition, this pattern was reversed, see Table 1.

Transcription and Coding

Interviews were videotaped and transcribed verbatim. For each of the two counterbalanced conditions (generative and inhibitory), there was one confirmation trial followed by two test trials. The following dependent measures were coded for each trial: what the child looked at first, what they explained first, their overall explanatory response, and the kind of responses or explanations they provided.

During the confirmation trial, the outcome the child looked at first was coded (starter/stopper =1, do-nothing=0), followed by the outcome the child mentioned first (starter/stopper=1, do-nothing=0, both=0.5). For example, if a child referred to both outcomes in their first explanation, it was coded as 0.5.

The two test trials per condition were counterbalanced; in half of the trials both of the boxes lit up first, and in half of the trials the boxes did not light up first. As in the confirmation trial, the outcome the child looked at first was coded (inconsistent=1, consistent=0). The outcome the child explained first was coded (inconsistent =1, consistent=0, both=0.5), followed by the explanations the child provided overall. Children's explanations for both inconsistent and consistent outcomes were coded separately. For example, if the children provided an explanation for the inconsistent outcome, "Cause this starter is broken, the parts inside don't work", and then went on to provide an explanation for the consistent outcome, "Cause this do-nothing is working, it doesn't make anything happen to the box", the child was coded as explaining the inconsistent outcome first (inconsistent =1, consistent = 0), and then their overall explanations were coded individually as providing an explanation for the inconsistent outcome (1=yes, 0=no) and providing an explanation for the consistent outcome (1=yes, 0=no).

Children were prompted several times for additional explanatory information and therefore had multiple opportunities to provide an explanation. For example, children were asked, “can you tell me more” or “do you have any other ideas” after each trial. The child was encouraged to provide as many explanations as possible. The trials were not terminated after the child’s first explanation, but only after the child indicated that they had no other information to share.

Children’s overall responses to the ambiguous causal questions were coded, and were not restricted to the child’s first explanation. Children’s explanations did not merely include the same kind of explanation with differing degrees of explanatory detail, they also provided different types of explanatory context. Overall explanations were coded into several kinds of causal and non-causal explanatory categories, for both consistent and inconsistent outcomes.

Explanatory Response Categories

The kind of explanation provided was coded for any outcome that the child mentioned (both consistent and inconsistent). Causal explanations for inconsistent outcomes were coded into 3 primary categories: category switch, causal function, and causal action.

In *category switch* explanations, children answered the explanation question by referring to a switch in category membership based on function. For example, “It is a do-nothing now” or “It’s really a starter, it only looks like a do-nothing” were coded as category switch explanations, see Table 2. Notice that to provide such explanations children had to ignore and go beyond perceptual identities and past history of perceptually identical objects.

Explanations that discussed a problem with the functioning of the object were coded as *causal function* explanations. These explanations included reference to the object being broken ("The stopper is not working anymore. It is broken"), the box being broken ("The box does not light anymore. It is broken"), or differences among the objects ("This one is heavier than the others"). Explanations that referred to insides or internal parts were also included in the causal function category ("There aren't stoppers inside" or "All out of batteries"). Explanations that referred to problems using the objects or problems with the placement of the objects were coded as *causal action* explanations ("You set the stopper/do nothing on the wrong sides" or "It's on the wrong box"), see Table 2.

Non-causal explanations for inconsistent outcomes were also coded into 3 primary categories: expectation violation, descriptive statements, and don't know. Explanations that described what could be expected to happen on the basis of appearance or past events without providing a cause were coded as *expectation violation* (e.g., "it wasn't supposed to turn on; I don't know why it did that"). Explanations that referred to the criteria children were using as a basis for their conclusion without further explanation of the cause were coded as *descriptive statements* (e.g., "It's not on because it's not glowing up"). If children were unable to provide an explanation or stated that they didn't know, their responses were coded as *don't know*. Explanations that referred to psychological constructs such as preference or desire (e.g., "maybe it wanted to") were also recorded.

Causal explanations for consistent outcomes were coded into three primary categories that were similar in kind to the explanatory categories for inconsistent outcomes: category label, causal function, and causal action. Explanations that referred to category membership using a label were coded as *category label* explanations. For example, “Because this one is a stopper” or “this one is a do-nothing” were coded as category label explanations. Note that unlike the category switch explanations for inconsistent outcomes, category label explanations for consistent outcomes involved providing an explanation based on the past history of perceptually identical objects.

Causal function explanations included reference to the object working or operating in some way (“Cause this one makes it work”), or features of the object (“This one is lighter”). Explanations that referred to insides or internal parts were also included in the function category (“There are batteries inside” or “electricity is in there”). Explanations that referred to placement of the objects were coded as *causal action* (“Cause that one’s on this box”).

Non-causal explanations for consistent outcomes were coded as descriptive statements. Explanations that referred to the criteria children were using as a basis for their conclusion without further explanation of the cause were coded as descriptive statements (e.g., “It’s on” or “It’s glowing”). If children were unable to provide an explanation or stated that they didn’t know, their responses were coded as *don’t know*. Note that unlike non-causal explanations for inconsistent outcomes children did not provide expectation-violation explanations for consistent outcomes.

Inter-rater reliability was established using a randomly selected sample of 25% of the data. Two persons independently coded the reliability sample with 95% agreement. Reliability was also calculated for explanatory coding categories for both consistent and inconsistent outcomes; with Kappas ranging from .86 to .94. All of the Kappas for this coding fall within near perfect (.80 and above) levels (Landis & Koch, 1977).

Results

In order to determine if events inconsistent with prior knowledge are especially likely to trigger causal explanation in young children (in contrast to events consistent with prior knowledge), the events that led to greater interest and attention, and most importantly, increased explanation, were examined. If the role of explanation is confirmatory, children should be interested in and provide explanations for the consistent event. If anomaly plays a special role in provoking children's explanations they should instead be specifically interested in and provide explanations for the inconsistent event (Figure 4). However, if state-change (boxes turning on or off) or negative outcomes (boxes off) are primarily responsible for triggering children's explanations I predict chance level performance given that in each test trial both contrastive outcomes included a state-change (both off to both on), negative outcome (both remain off), neither (both remain on), or both (both on to both off).

Which outcomes do children look at and explain first?

Test trials. Data from Study 1 indicate that events that were inconsistent with children's prior knowledge were very likely both to attract children's attention and also to provoke children's explanations. Thus, across age groups and test trials (out of 4 possible trials), children were much more likely to look first at the inconsistent outcomes than the consistent outcomes, (M=3.73), $t(47) = 18.61$, $p < .001$. Likewise, children were much more likely to explain the inconsistent outcome first, (M=3.18), $t(47) = 11.72$, $p < .001$, Figure 5. An overall chi-square analysis comparing children that looked at the inconsistent outcome first 0-2 versus 3-4 times (out of a total of 4 test trials) demonstrates that the majority of children (90%) looked at the inconsistent outcome first in 3 or more of the 4 test trials, X^2 (1, N = 48) = 30.08, p<.001. The same analysis comparing children that explained the inconsistent outcome first 0-2 versus 3-4 times demonstrated the same pattern of results; the majority of children (77%) explained the inconsistent outcome first in 3 or more of the 4 test trials, X^2 (1, N = 48) = 14.08, p<.001.

In order to investigate potential age and condition-related effects on the outcomes children looked at first and explained first, two separate age group (3-, 4-, 5-year-olds) X condition (inhibitory versus generative) X final-outcome (both boxes on versus both boxes off) repeated measures ANOVAs were conducted. Condition and final-outcome were within-subjects factors and age group was a between-subjects factor. Due to the fact that across age groups and conditions children were very likely to look at and explain the inconsistent outcome first, there was no significant main effect of age (for either the first look or first explanation analysis). In addition there was no significant main effect of condition (inhibitory versus generative) or final-outcome (both on versus both off). However, for the first explanation analysis, the condition X final-outcome interaction was significant, $F(1,45) = 4.62$, $p < .05$. Posthoc tests indicate that for the inhibitory condition, children were more likely to explain the inconsistent outcome when both boxes remained on (no state change) (M=.84 out of 1.0) than when both boxes turned off (state change) (M=.70), $p < .05$. One possible explanation for this finding is that children may have found it more surprising that an object they anticipated would activate the box (stopper) failed to do so than when an object they anticipated would have no effect on the box (do-nothing) instead acquired a new functional property (stopping the box). This would be consistent with the finding in the confirmation trials that children are more likely to explain the object that changed the state of the box (activation or deactivation) than the object that had no effect on the box. Importantly however, in both cases they were significantly more likely to look at the inconsistent outcome first than chance alone would predict, *p*s<.001.

Confirmation trials. As a baseline comparison to children's test-trial explanations, the events children looked at first, in addition to the events children explained first were analyzed for the confirmation trials. In the confirmation trials, children were asked to explain an initial paired event in which both objects of each kind functioned in a manner consistent with prior knowledge. Out of 2 possible trials, children were more likely both to look first at the object that changed the state of the box (starter or stopper), (M=1.26), $t(47) = 2.84$, $p < .01$, and to first explain the object that changed the state of the box, (M=1.41), $t(47) = 4.57$, $p < .001$, as compared to the object that had no effect on the box (do-nothing). Children were significantly more likely to first explain the object that changed the state of the box in both the generative (M=.69) and the inhibitory (M=.74) conditions, *p*s<.001 (comparisons to chance). This finding indicates that in the absence of an inconsistent event or outcome, state-change serves as a trigger for children's explanations.

Which outcomes do children provide an explanation for overall?

Given that children were prompted several times for additional explanatory information and therefore had multiple opportunities to provide explanations, it is also revealing to analyze whether, overall, children are more likely to explain events consistent or inconsistent with their prior knowledge. Note that the entirety of children's explanations for both inconsistent and consistent events were coded and were not in fact restricted to their first explanation. Out of a total of 4 possible trials, providing an explanation for inconsistent outcomes (M=3.98), was much more likely than providing an explanation for the consistent outcomes (M=2.73), $t(47) = 6.38$, $p < .001$ (Figure 6).

In order to investigate potential age, condition-, and final outcome-related effects on which events children provided an explanation for overall, an age group (3-, 4-, 5-year-olds) X condition (inhibitory versus generative) X final-outcome (both boxes on versus both boxes off) X outcome-explained (inconsistent versus consistent outcome) repeated measures ANOVA was conducted. Condition, final-outcome, and outcome-explained were within-subjects factors and age group was a between-subjects factor. There were significant main effects for age group $F(2,45) = 14.18$, $p < .001$, condition, $F(1,45) = 12.71$, $p < .001$, and, most importantly, outcome explained $F(1,45) = 63.12$, $p < .001$. Posthoc tests indicate that overall, 4- and 5-year-olds (M=.89, M=.94 out of 1.0, respectively) were more likely to provide explanations than 3-year-olds (M=.69), and explanations were more prevalent in the inhibitory (M=.88) than the generative condition (M=.80). One possible explanation for this effect is that children may find the initial-state of the light boxes being (inhibitory condition) more novel or unexpected than the initial-state of the light boxes being off (generative condition) and may therefore feel more compelled to provide explanations overall in the inhibitory condition.

The significant main effect for outcome explained points to the centrality of the a priori contrast between inconsistent and consistent outcomes. Therefore the next set of analyses focus more precisely on that comparison. The outcome explained X age group interaction was significant, $F(2,45) = 13.90$, $p < .001$, indicating developmental differences in the outcome children were most likely to provide an explanation for overall. Posthoc tests reveal that although 3- and 4-year-olds were more likely to explain inconsistent versus consistent outcomes, *p*s < .001, 5-year-olds were equally likely to explain both overall (Figure 7a). This finding may indicate that older children may be more likely to provide explanations for all available outcomes and evidence than younger children, especially when they are given multiple opportunities for doing so. The condition X outcome explained interaction, $F(1,45) = 12.62$, $p < .001$, the condition X final-outcome interaction, $F(1,45) = 6.88$, $p < .05$, and the three-way interaction with condition X final-outcome X outcome explained, $F(1,45) = 8.25$, $p < .001$, were all significant. Posthoc tests reveal that although children were more likely to provide explanations for consistent outcomes in the inhibitory (M=.76) than the generative condition (M=.60), $p < .001$, this was most likely when both boxes were off (state-change), p<.001. However, for all four test trials (in each condition and final-outcome) children were more likely to provide an explanation for an inconsistent than a consistent outcome.

Individuals' response patterns provide an important complementary analysis. For this analysis, the dependent variables were whether children provided an explanation for inconsistent outcomes and whether children provided an explanation for consistent outcomes. If children provided an explanation 0-2 times they were given a score of 0, and if they provided an explanation 3-4 times (out of a total of 4 test trials), they were given a score of 1 (for each outcome). Individuals' patterns of responses confirm that children were more likely to provide an explanation for inconsistent outcomes than consistent outcomes: although 29 of 48 children (60%) consistently provided explanations for both inconsistent and consistent events across 3 or more test trials, the remaining children (19 out of 48 children, or 40%) explained inconsistent events on 3 or more test trials and explained consistent events on 2 or fewer test trials. Notably, the reverse was never true; none of the children explained consistent events on 3 or more test trials while explaining inconsistent events on 2 or fewer test trials, McNemar's $\chi^2(1) = 17.05$, $p < .001$.

Comparing the kinds of causal explanations children provided for consistent and inconsistent outcomes

In addition to analyzing which kind of event children explained first and explained overall, it is also informative to analyze the kinds of overall explanations children provided for consistent and inconsistent outcomes. Children's explanations for both inconsistent and consistent outcomes were coded into three primary causal categories: causal category, causal function, and causal action. The majority of children's explanations at all age groups were causal, for both inconsistent outcomes (55% of the explanations provided by three-year-olds, 81% of the explanations provided by 4-year-olds, and 95% of the explanations provided by 5-year-olds) and consistent outcomes (76% of the explanations provided by three-year-olds, 74% of the explanations provided by 4-year-olds, and 88% of the explanations provided by 5-year-olds).

An age group (3-, 4-, 5-year-olds) X condition (inhibitory X generative) outcome explained (consistent versus inconsistent) X causal explanation type (causal function, causal action, category label/category switch) repeated measures ANOVA was conducted in order to investigate age and condition-related changes in the kinds of causal explanations children provided. Condition, outcome explained and causal explanation type were within-subject factors and age group was the between-subjects factor. Causal explanation type was the dependent variable.

There was a significant main effect for age group, $F(2,45) = 10.83$, $p < .001$, indicating that there was a significant increase in causal explanations for both consistent and inconsistent outcomes with age. Posthoc tests indicate that 4- and 5-year-olds were more likely to provide causal explanations than 3-year-olds, *p*s<.05. Most importantly, there was also a significant main effect for outcome explained, $F(1,45) = 24.19$, $p < .001$ indicating that children were significantly more likely to provide causal explanations for inconsistent outcomes (M=3.98) than consistent outcomes, (M=2.73).

The main effect for causal explanation type was also significant, $F(2,90) = 11.64$, $p < .001$. Causal function and causal category explanations were most common (and not significantly different from one another), and each was more frequent than causal action explanations, *p*s< .01, (Figure 7). Additionally, the three-way age group X outcome explained X causal explanation type interaction was significant, $F(4,84) = 2.74$, $p < .05$. Posthoc tests revealed that 5-year-olds were more likely to give causal category explanations for inconsistent outcomes (category switch) (M=2.13 out of 4) than causal category explanations for the consistent outcomes (category label) (M=1.25), p<.001. Explanations that made use of psychological explanations were very rare; only one 3-year-old and two 4-year-olds used psychological language in their explanations ("This one thinks it's a starter").

In supplementary analyses of children's explanations for inconsistent outcomes only, whereas causal function and causal action explanations remained constant across age groups (M=1.43 out of 4 total explanations), $F(2,47) = .18$, *ns*, category switch explanations increased with age (16% of total explanations for 3-year-olds, 38% of total explanations for 4-year-olds, and 53% of total explanations for 5-year-olds were category switch explanations), $F(2,47) = 3.95$, $p < .05$ (Figure 8a). The finding of category switch explanations in children so young is especially noteworthy given that an explanation of this kind required overriding perceptual appearances and prior knowledge about an object and spontaneously redefining category boundaries around function.

For consistent outcomes, whereas causal function explanations (M=1.19 out of 4 total explanations) and causal action explanations (M=0.21) remained constant across age groups, $F(2,47) = 1.47$, *ns*; $F(2,47) = 1.78$, *ns* (respectively), category label explanations increased with age (6% of total explanations for 3-year-olds, 24% of total explanations for 4-year-olds, and 31% of total explanations for 5-year-olds were category label explanations, M=0.81), $F(2,47) = 3.56$, $p < .05$. The increase in category label explanations with age for consistent outcomes parallels the increase in category switch explanations with age for inconsistent outcomes.

In addition to analyzing the proportion of total explanations that fell into each explanatory category, the percentage of participants that gave category switch, causal function, and causal action explanations at least once was calculated. This analysis is a useful supplement to the data on proportion of total explanations of each explanatory category because it provides evidence that although 5-year-olds are more likely to provide category switch explanations than younger children, overall, a substantial portion of 3- and 4-year-olds provided this kind of explanation at least once (Table 4a). Additionally, a substantial portion of children provided more than one kind of causal explanation for both inconsistent (Table 5a) and consistent (Table 5b) outcomes.

Discussion

The objective of Study 1 was to investigate two competing hypotheses about the function of children's explanations. If the role of explanations for children is confirmatory, children should be interested in and provide explanations for the consistent event. If "anomaly" plays a special role they could be specifically interested in and provide hypotheses and explanations for the inconsistent event. A third possibility is that what is relevant is neither consistency not inconsistency, but other factors (such as state-change or negative outcomes). In order to investigate this experimentally, events that were inconsistent with children's prior knowledge were simultaneously contrasted with events that were consistent with children's prior knowledge. State-change and the negative versus positive outcomes were controlled for by being held constant. The outcomes which led to greater interest and attention, and, most importantly, increased explanation were of primary interest. The outcome children looked at first, explained first, and were more likely to provide an explanation for overall were the primary dependent measures. Moreover, the content of children's explanations for both inconsistent and consistent outcomes was examined, specifically whether children would provide explanations primarily in terms of surface features and past histories, or whether they would offer explanations focused on less-obvious properties and underlying causal properties.

The data from Study 1 provide support for the thesis that inconsistent outcomes are an especially powerful trigger for children's explanations. Across the generative and inhibitory conditions, children were much more likely to attend to and explain the outcome that was inconsistent with their prior knowledge. This is significant for several reasons. If children had just been paying attention to the perceptual outcome (state-change or negative outcome), the event they explained should have been at chance, given that in the test trials the outcomes were identical (both boxes were on or off). Indeed, data from the confirmation trials provide support for the hypothesis that in the absence of an inconsistent event or outcome, state-change serves as a trigger for children's explanations. Additionally, one of the strengths of the design is that children were not primed to provide an explanation for the inconsistent outcome prior to the test trials. In the confirmation trial, children were asked an ambiguous causal questions referring to either visible outcome, both of which were consistent with children's prior knowledge about how the objects function. Finally, given that children already had an explanation for the consistent outcome, they could have easily ignored the inconsistent outcome (for which they were given no explanation) and just explained the consistent outcome. Instead the data indicate that children find outcomes inconsistent with prior knowledge especially noteworthy, and are not only more likely to look at and explain an inconsistent outcome first, they are also more likely to provide an explanation for an inconsistent outcome overall.

However, it is not the case that children exclusively provide explanations for inconsistent outcomes and therefore that children's explanations are never confirmatory. Instead, the data from Study 1 demonstrate that children's explanations were *not merely confirmatory*. In fact, constructing explanations for events that are consistent with children's prior knowledge and experience may indeed be an important function of children's own explanations. For example, explaining consistent outcomes may provide children with an important opportunity to deepen their understanding of causal phenomena by allowing children to generate causal mechanisms. Importantly, the thesis is that outcomes inconsistent with prior knowledge are more likely to motivate causal explanatory cognition, not that children do not or cannot generate explanations for outcomes consistent with their prior knowledge.

How children explain events provides insight into their understanding of the causal basis for the outcome. Therefore, the kinds of explanations children provided for both the inconsistent and the consistent outcomes were of interest. Children provided three distinct kinds of causal explanations for both inconsistent and consistent outcomes, namely, causal category explanations, causal function explanations, and causal action explanations. Causal category explanations for inconsistent outcomes referred to a *category switch*. For example, "it's really a starter; it only looks like a do-nothing". Explanations of this kind are especially noteworthy because they privilege functional information by referring to underlying, internal causal properties, overriding perceptual appearances. Children also provided causal category explanations for consistent outcomes by using the object's label to refer to its kind, "Because it's a starter. Starters make the light in the boxes turn on". However, it is worth pointing out that category switch explanations for inconsistent outcomes are quite different from causal label explanations for consistent outcomes. In the case of a category switch explanation for inconsistent outcomes, children spontaneously re-defined category boundaries around function by providing a label for the object that was shared by items that were functionally similar and perceptually dissimilar. In contrast, causal label explanations referred to category membership with objects similar in both functional history and perceptual appearance. Additionally, causal category explanations were most common among 5-year-olds, and 5-year-olds were more likely to provide causal category explanations for inconsistent outcomes than consistent outcomes.

Across age groups, causal function explanations were given quite frequently overall, for both inconsistent and consistent outcomes. Causal function explanations for inconsistent outcomes referred to alternative or problematic functioning, and for consistent outcomes referred to the object functioning or operating correctly. Interestingly, causal function explanations for both kinds of outcomes often included information about potential inside parts (power, batteries) or underlying causal properties (energy, electricity). Finally, causal action explanations referred to the placement or use of the objects, although were less frequent than either causal category or causal function explanations overall.

The content of children's explanations for the outcomes in this study is revealing of the nature of children's object concepts. Although it is well documented that even 3-year-olds can categorize objects in terms of novel, non-obvious properties (Gelman & Coley, 1990; Gelman & Markman, 1986, 1987), apply names to objects with the same functional properties (Kemler-Nelson, 1995), and categorize and name objects based on novel causal properties (Gopnik & Sobel, 2000), it is not known whether and how preschool children might make use of this information in their causal explanations. I argue that children's explanations may be especially revealing of their causal understanding; therefore the finding that children's explanations typically include information about function and causal properties is a strong indication that children use information about causal properties and function when reasoning about objects.

However, several open questions remain concerning the use of functional labels in Study 1. Many of children's explanations referred to the functioning of the objects or light boxes in some fashion. This was expected, but additionally some responses might have been especially scaffolded and provoked in Study 1 because the object names (starters, stoppers, and do-nothings) were straight-forward descriptions of their functions. Were children's explanations for inconsistent events motivated primarily by the use of the labels? Were children really surprised at the inconsistent outcome or surprised by the inconsistency of the function with the label?

Additionally, one of the most interesting kinds of explanations provided by children in Study 1 were category switch explanations. However, given the nature of the functional label, it is unclear whether children actually relabeling around new category boundaries or whether they were simply giving a more literal description instead. Study 2 was designed specially to address these alternative possibilities.

Chapter III

Study 2: Investigating Outcomes Inconsistent with Prior Knowledge as a Trigger for Causal Explanatory Reasoning in Young Children Using Novel Labels

The objective of Study 2 was to investigate whether explanations for events not in accord with prior knowledge could have been scaffolded by the use of functional labels (“starters”, “stoppers”, and “do-nothings”) in Study 1, specifically because the objects violated their labeled functions. Children’s interest in and explanations for inconsistent outcomes in Study 1 were informative nonetheless; however, it is important to consider the possibility that children’s attention and explanations were driven by the fact that the events they observed contradicted the objects’ labeled function, instead of inconsistency with the anticipated outcome. In order to investigate this possibility and to determine if the same effects are found without this cue, Study 2 was designed to examine the effect of novel, non-functional labels for the objects (toma, blicket) and their influence on (a) the kinds of outcomes (consistent or inconsistent) children attend to and explain and (b) the kinds of explanations children provide for these outcomes.

Given evidence that even 3-year-olds can extend novel labels to new categories based on function (Gelman & Markman, 1986, 1987), I predicted that information about function and underlying causal properties would be found in children’s causal explanations, even when the labels themselves are novel and do not, in themselves provide additional causal information. Furthermore, the data from Study 1 demonstrate that by redefining category membership around shared function instead of perceptual appearances and invoking underlying, internal causal properties and mechanisms, children's causal explanations may be especially revealing of the nature of children's object concepts. Additionally, I predicted that children would focus more on explaining events inconsistent with prior knowledge, even when the labels themselves do not carry such functional information.

Method

Participants

Sixteen 3-year-olds (*M* age 3,5; range 3,0 to 3, 10) and sixteen 4-year-olds (*M* age 4,4; range 4, 0 to 4, 11) were recruited from a Midwestern university town, and were primarily White. Approximately equal numbers of boys and girls participated in the study. None of the participants in Study 2 participated in Study 1.

Materials

The materials used in Study 2 were identical to those from Study 1.

Procedure

The procedure for Study 2 was modeled on Study 1, with a critical modification: the objects were given non-functional, novel labels ("a toma" / "not a toma" and "a blicket" / "not a blicket"), and were never referred to as "a starter", "a stopper", or "a do-nothing". The confirmation and test trials were identical to those of Study 1.

Transcribing and Coding

The same procedure for coding Study 1 was used to code Study 2. Table 3 provides examples of category switch, causal function, and causal action explanations from Study 2. Inter-rater reliability was established using a randomly selected sample of 25% of the data. Two persons independently coded the reliability sample with 98% agreement. Reliability was calculated for explanatory coding categories for both consistent and inconsistent outcomes with Kappas ranging from .92 to .94. All of the Kappas for this coding fall within near perfect (.80 and above) levels (Landis & Koch, 1977).

Results

In order to determine if events inconsistent with prior knowledge are especially likely to trigger causal explanation in young children (in contrast to events consistent with prior knowledge), the events that led to greater interest and attention, and most importantly, increased explanation, were examined. If the role of explanation is confirmatory, children should be interested in and provide explanations for the consistent event. If anomaly plays a special role in provoking children's explanations they should instead be specifically interested in and provide explanations for the inconsistent event (Figure 4).

Which outcomes do children look at and explain first?

Test trials. As in Study 1, inconsistent outcomes not only attracted children's attention but also provoked children's explanations. Across age groups, children were much more likely to look first at the inconsistent outcomes than the consistent outcomes (M=3.25 out of 4), $t(31) = 7.44$, $p < .001$. Likewise, children were much more likely to explain the inconsistent outcome first (M=3.05), $t(31) = 9.86$, $p < .001$ (Figure 5).

An overall chi-square analysis comparing children who looked at the inconsistent outcome first 0-2 versus 3-4 times (out of a total of 4 test trials) demonstrates that the majority of children (78%) looked at the inconsistent outcome first in 3 or more of the 4 test trials, X^2 (1, N = 32) = 10.13, p<.001. The same analysis comparing children who explained the inconsistent outcome first 0-2 versus 3-4 times demonstrated the same pattern of results; the majority of children (75%) explained the inconsistent outcome first in 3 or more of the 4 test trials, X^2 (1, N = 32) = 8.00, p<.01.

In order to investigate potential age and condition-related effects on the outcomes children looked at first and explained first, an age group (3-, 4-year-olds) X condition (inhibitory versus generative) X final-outcome (both boxes on versus both boxes off) repeated measures ANOVA was conducted. Condition and final-outcome were within-subjects factors and age group was a between-subjects factor. Due to the fact that across age groups and condition children were very likely to look at the inconsistent outcome first, and explain the inconsistent outcome first, there were no significant effects (for either first look or first explanation). Unlike Study 1, none of the interactions were significant.

Confirmation trials. As a baseline comparison to children's test-trial explanations, the events children looked at first, in addition to the events children explained first were analyzed for the confirmation trials. In the confirmation trials (scores ranging from 0-2), children were more likely to look first at the object that changed the state of the box (started or stopped it) than at the object that did not change the state of the box (M=1.28 out of 2), $t(31) = 2.06$, $p < .05$, and more likely to explain the object that changed the state of the box than the object that had no effect on the box

(*M*=1.23), *t*(31) = 1.97, *p*<.05. However, although children were significantly more likely to first explain the object that changed the state of the box in the generative (*M*=.70), *t*(31) = 2.63, *p* < .01, they were equally likely to explain the object that changed the state of the box as the object that had no effect on the box in the inhibitory (*M*=.53) condition, *t*(31) = .32, *ns* (comparisons to chance). One possible explanation for this is that children may be more compelled by an object that fails to turn the light box off (thereby keeping the light box on) than an object that fails to turn the light box on (thereby keeping the light box off).

Which outcomes do children provide an explanation for overall?

As in Study 1, children were much more likely to explain events inconsistent with their prior knowledge than events consistent with their prior knowledge. Out of a total of 4 possible trials, providing an explanation for the inconsistent outcome (*M*=3.94) was more likely than providing an explanation for the consistent outcome (*M*=2.78), *t*(31) = 4.41, *p* < .001 (Figure 6).

In order to investigate potential age and condition-related effects on which outcomes children provided an explanation for in the test trials, an age group (3-, 4-year-olds) X condition (inhibitory versus generative) X final-outcome (both boxes on versus both boxes off) X outcome-explained (inconsistent versus consistent outcome) repeated measures ANOVA was conducted. Condition, final-outcome, and outcome-explained were within-subjects factors and age group was a between-subjects factor. The only significant effect was a main effect for outcome explained, *F*(1,30) = 22.64, *p* < .001. This confirms the results from Study 1 that children are much more likely to provide explanations for inconsistent versus consistent outcomes. Both 3-year-olds (*t*(15) = 3.47,

$p < .01$) and 4-year-olds ($t(15) = 2.7$, $p < .05$) were significantly more likely to explain inconsistent than consistent outcomes (Figure 7b). The lack of significant effects for age group and condition in Study 2 may also reflect the more narrow age range in Study 2 (only 3- and 4-year-olds).

Individuals' response patterns provide an important complementary analysis. For this analysis, the dependent variables were whether children provided an explanation for inconsistent outcomes and whether children provided an explanation for consistent outcomes. If children provided an explanation 0-2 times they were given a score of 0, and if they provided an explanation 3-4 times (out of a total of 4 test trials), they were given a score of 1 (for each outcome). Individuals' patterns of responses confirm that children were more likely to provide an explanation for inconsistent outcomes than consistent outcomes: although 19 of 32 children (59%) consistently provided explanations for both inconsistent and consistent events across 3 or more test trials, 13 out of 32 children (41), explained inconsistent events on 3 or more test trials and explained consistent events on 2 or fewer test trials. Notably, the reverse was never true; none of the children explained consistent events on 3 or more test trials while explaining inconsistent events on 2 or fewer test trials, McNemar's $\chi^2(1)= 11.08$, $p < .001$.

Comparing the kinds of causal explanations children provided for consistent and inconsistent outcomes

As in Study 1, the majority of children's explanations at all age groups were causal, for both inconsistent outcomes (81% of the explanations provided by three-year-olds, 84% of the explanations provided by 4-year-olds) and consistent outcomes (63% of the explanations provided by three-year-olds, 74% of the explanations provided by 4-

year-olds). An age group (3-, 4-year-olds) X condition (inhibitory versus generative) X outcome explained (consistent versus inconsistent) X causal explanation type (causal function, causal action, category label/category switch) repeated measures ANOVA was conducted in order to investigate age-related changes in the kinds of causal explanations children provided. Condition, outcome explained and causal explanation type were within-subject factors and age group was the between-subjects factor. Causal explanation type was the dependent variable.

As predicted, there was a significant main effect for outcome explained, $F(1,30) = 25.94$, $p < .001$ indicating that children were significantly more likely to provide causal explanations for inconsistent outcomes (M=3.94) than consistent outcomes (M=2.78). Unlike Study 1, the main effect for age group was not significant, $F(1,30) = 21.34$, *ns*, indicating that 3-year-olds were just as likely as 4-year-olds to provide causal explanations.

The main effect for causal explanation type was also significant, $F(2, 60) = 9.77$, $p < .001$. Posthoc tests indicate that although children were equally likely to provide causal function and causal category explanations overall, *ns,* and equally likely to provide causal category and causal action explanations overall, *ns*, they were more likely to provide causal function explanations than causal action explanations, $ps < .01$ (Figure 8b). The three-way age group X outcome explained X causal explanation type interaction was also significant, $F(2, 60) = 5.64$, $p < .01$. Posthoc tests revealed that 3-year-olds were more likely to give causal function explanations for inconsistent (M=2.5 out of 4) than consistent outcomes (M=1.4), $p<.01$. Four-year-olds were more likely to give causal action explanations for inconsistent (M=.5) than consistent outcomes, (M=.12) $p<.05$, and more likely to give category switch explanations for inconsistent outcomes (M=1.9) than category label explanations for the consistent outcomes (M=.94), p<.01. Explanations that made use of psychological causes/descriptions were very rare; only two 3-year-olds and one 4-year-old used psychological language in their explanations (“This one doesn’t know how to act, it isn’t working right”).

Additionally, as in Study 1, there were changes with age in the kinds of causal explanations that children provided for inconsistent outcomes. Whereas causal action explanations remained constant across age groups (M=.38 out of 4), $F(1,31) = .52$, *ns*, causal function explanations decreased with age, $F(1,31) = 13.04$, $p < .001$ (63% of total explanations for 3-year-olds, and 24% of total explanations for 4-year-olds were causal function explanations). As in Study 1, category switch explanations increased with age, (13% of total explanations for 3-year-olds and 47% of total explanations for 4-year-olds were category switch explanations), $F(1,31) = 7.88$, $p < .01$ (Figure 8b).

Whereas causal function explanations (M=1.38 out of 4 total explanations) and causal action explanations (M=.06) for consistent outcomes remained constant across age groups, $F(1,31) = .06$, *ns*; $F(1,31) = .4$, *ns* (respectively), category label explanations increased with age (5% of total explanations for 3-year-olds and 24% of total explanations for 4-year-olds were category label explanations, M=.56 out of 4 possible explanations), $F(1,31) = 4.92$, $p < .05$. As in Study 1, the increase in category label explanations with age for consistent outcomes parallels the increase in category switch explanations with age for inconsistent outcomes.

In addition to analyzing the proportion of total explanations that fell into each explanatory category, the percentage of participants that gave category switch, causal function, and causal action explanations for inconsistent outcomes at least once was calculated. These results are informative because, as in Study 1, they indicate that a substantial portion of 3-year-olds and the majority of 4-year-olds provided a category switch explanation, but also that 3-year-olds were very likely to provide explanations based on the causal functioning of the objects, even though the objects were given novel, non-functional labels (Table 4b). Furthermore, a substantial portion of children provided more than one kind of causal explanation for both inconsistent (Table 5a) and consistent (Table 5b) outcomes.

Comparing the results of Studies 1 and 2

Across Studies 1 and 2, inconsistent outcomes were very likely to provoke children's explanations. Although children were more likely to look at the inconsistent outcome first in Study 1 than in Study 2, $F(1,79) = 7.24$, $p < .01$, children were equally likely to explain the inconsistent outcome first, $F(1,79) = .75$, *ns*. The similarity of the results across studies is important because they provide confirmatory evidence that (a) inconsistency with prior knowledge is a powerful trigger for children's causal explanations and (b) the pattern of data from Study 1 cannot merely be attributed to the use of functional labels.

Overall, the results from the confirmation trials in Studies 1 and 2 were highly similar, consistent with the possibility that in the absence of information about inconsistency with prior knowledge, children are more likely to explain a state-change outcome than an outcome that did not change state. However, one difference between the two studies was that in Study 1 children were more likely to explain the object that changed the state of the light box first in both conditions, whereas in Study 2, children were only more likely to do so in the generative condition. One potential explanation for this is that children may have been equally compelled to first explain an object that maintained a positive outcome (keeping the light on) as an object that both changed the state of the box and produced a negative outcome. Although this result is not consistent with my predictions, it may be that the outcome was not perceived as sufficiently "negative" to provoke children's explanations. Another possibility is that the smaller sample size participating in Study 2 (32 instead of the 48 children who participated in Study 1) may have contributed to this result.

I also compared the kinds of causal explanations children provided in Studies 1 and 2. In order to investigate whether children would provide different kinds of explanations when functional labels (Study 1) versus novel labels (Study 2) were used and whether the use of different labels would influence the kind of outcome children explained, a study (Study 1 versus Study 2) X age group (3-, 4-, 5-year-olds) X condition (inhibitory versus generative) X outcome explained (consistent versus inconsistent) X causal explanation type (causal function, causal action, category label/category switch) repeated measures ANOVA was conducted. Outcome explained and causal explanation type were within-subject factors, and age group and condition were between-subjects factors. Causal explanation type was the dependent variable. Because the results for Studies 1 and 2 have already been described separately, only significant results with study as a factor will be included here.

The only significant interaction was the four-way study X age group X outcome explained X causal explanation type, $F(2, 140) = 4.35$, $p < .05$. Posthoc tests revealed that the only significant difference between studies is that 4-year-olds were more likely to give causal action explanations for inconsistent outcomes in Study 2 than in Study 1, $p<.05$. However, given that the overall rate of causal action explanations across both studies was very low (less than 10% of total explanations), this finding is unlikely to represent a substantive difference in the content of children's causal explanations across studies.

Overall, Study 2 replicates all of the core findings of Study 1, providing additional support for the thesis that outcomes inconsistent with prior knowledge are especially powerful explanatory triggers. Across Studies 1 and 2, children of all ages were more likely to look at and explain inconsistent outcomes first on the test trials. Additionally, overall, children were more likely to provide an explanation for an inconsistent outcome than a consistent outcome, even with multiple opportunities to explain both. Finally, the content of their explanations, specifically causal category and causal function explanations, referred to underlying, internal causal mechanisms, overriding perceptual appearances.

Discussion

Children's causal explanations are especially revealing of the nature of children's object concepts by (a) privileging function information over perceptual features, (b) evoking causal mechanisms and internal, underlying causal properties, and (c) spontaneously using novel labels to redefine category boundaries around shared function. The data from Study 2 demonstrate that, as in Study 1, children focused more on explaining events inconsistent with prior knowledge than events consistent with prior knowledge, even when the labels themselves do not provide additional functional information. Further, information about function and underlying causal properties characterized children's causal explanations, even when the labels for the objects were novel and did not provide additional causal information.

Study 2 provides further support for the hypothesis that inconsistent outcomes are an especially powerful trigger for children's explanations. As in Study 1, across the generative and inhibitory conditions, children were much more likely to attend to and explain the inconsistent outcome, overriding perceptual appearances in favor of function information. Importantly, the data demonstrate that the category switch explanations in Study 1 were not merely redescription. Even children as young as three years old use information about underlying causal properties to re-define category membership in their explanations, even when the labels are novel and not tied to function. Additionally, children's explanations referring to causal function cannot be attributed to restating functional information entailed in the label, as they did not have access to this information in Study 2. Finally, Study 2 provides evidence that children were not simply drawn to explain events based on the inconsistency of the outcome with the label, but instead were motivated to provide explanations for outcomes that were inconsistent with their prior knowledge.

Discussion of Studies 1 and 2

The primary objective of Studies 1 and 2 was to address two distinct hypotheses about the function of children's own explanations; explanation as confirmation or explanation as discovery. In order to investigate these hypotheses experimentally, in the test trials, events that were consistent with children's prior knowledge were contrasted with events that were inconsistent with prior knowledge and children were asked to provide an explanation.

Data from Studies 1 and 2 demonstrate that inconsistent outcomes are an especially powerful trigger for children's explanations. Across the generative and inhibitory conditions, children were much more likely to attend to and explain the inconsistent outcome. Notably, this result provides firm evidence that young children's explanations do not serve an exclusively confirmatory role. Children could easily have provided more explanations for the outcome they already had an explanation for (the consistent outcome) and ignored the inconsistent outcome (for which they were given no explanation). However, this was not the case. My data are thus consistent with the idea that explanation provides children with the opportunity to articulate new hypotheses for events that, at first, disconfirm their current knowledge.

Children's explanations for inconsistent outcomes referred to underlying, internal causal properties, overriding perceptual appearances. For example, children explained inconsistent outcomes by stating that, "all the energy is gone" or "maybe that toma doesn't have batteries in it". These data suggest that children are looking for explanations that extend beyond the available evidence. Furthermore, children's explanations may play an active, important role in the learning process. If children use explanation as a mechanism for acquiring new knowledge, they should provide explanations for events that have the potential to teach them something new. Explaining inconsistent events provides just such an opportunity, and such explanatory biases might aid in learning. Although my data do not speak to this question directly, an interesting direction for future research would be to directly examine the influence of children's explanations on learning and knowledge acquisition.

Because the primary motivation for Studies 1 and 2 was to manipulate prior knowledge experimentally, I chose to construct a task that would allow for precise control over participants' background knowledge about specified physical outcomes. Inconsistency with prior knowledge was of primary interest, and disentangling inconsistency with prior knowledge from other potential explanatory triggers was a crucial component of the study design. However, a noteworthy feature of the design of Studies 1 and 2 was to both control for additional potential explanatory triggers (state-change and negative outcomes) by holding them constant, and to investigate state-change as a potential explanatory trigger in the absence of an inconsistent outcome (confirmation trials). For example, state-change was controlled for in the test trials of Studies 1 and 2 by both outcomes changing equally (both turned off/ turned on). The valence of the outcome was controlled for in the test trials by the outcomes being identical (both off/on). Thus for all of the test trials, the only difference between the outcomes was whether the outcome was consistent or inconsistent with prior knowledge, which was carefully controlled by the experimenter in the first part of the tasks.

It is possible that as distinct kinds of inconsistent outcomes, state-change and negative outcomes may all provoke children's explanations. Because these kinds of outcomes are often confounded both in the world and in experimental manipulations, an overarching objective of Study 3 was to begin to experimentally differentiate the kinds of events that children find noteworthy and therefore feel compelled to explain. In order to do so, Study 3 focused specifically on state-change and negative outcomes.

Chapter IV

Study 3: Investigating State-change and Negative Outcomes as Triggers for Causal Explanatory Reasoning: Children's Explanations of Curing and Preventing Illness

Study 3 expands upon Studies 1 and 2 in three ways by (a) extending the findings from Studies 1 and 2 to a new domain, (b) specifically targeting state-change and negative outcomes as additional kinds of explanatory triggers, and (c) investigating causal explanatory triggers in a knowledge-rich context. I chose to investigate causal explanatory triggers in the domain of illness for several reasons. The first is that developmental research in the domain of folk biology has demonstrated that young children have complex and often elaborate beliefs about biological processes at an early age (Inagaki & Hatano, 2002; Wellman & Gelman, 1992). Even 3-year-olds can recognize biological causes (Kalish, 1996) and can explicitly use appropriate language to articulate this understanding (Wellman, Hickling, & Schult, 1997). This indicates that preschool children have the capacity to generate biological explanations and that the biological domain is an appropriate place to investigate the kinds of events and outcomes that trigger causal explanatory reasoning in children.

Children are also very interested in biological phenomena and even 3-year-olds can provide biological explanations that refer to invisible, non-obvious processes and entities (Au, Sidle, & Rollins, 1993; Legare et al., under review). Finally, state-change and negative valence outcomes are both meaningful and relevant in the domain of illness. Studying state-change and negative outcomes that involve the physical effects of objects on light boxes (as in Studies 1 and 2), may be less informative than studying state-change and negative outcomes in a domain in which the outcomes are more consequential. For example, it is possible that a box turning from on to off is not sufficiently negative or noteworthy (as an isolated outcome) to provoke children's causal explanations. However, explaining why a healthy person got sick, or a sick person failed to recover from an illness may be substantially more engaging and therefore may be more likely to motivate children to construct a causal explanation. Study 3 was designed to investigate this possibility directly, using methodology that is analogous, but distinct from the contrastive outcome design used in Studies 1 and 2.

Because the primary objective of Studies 1 and 2 was to investigate and isolate outcomes inconsistent with prior knowledge as a potential trigger for children's causal explanations, children's prior knowledge was experimentally controlled. However, in Study 3, the objective was to focus instead on a domain in which young children already have a substantial amount of prior knowledge. Therefore it wasn't possible to control for or quantify prior knowledge in the same way. Nonetheless, children's explanations for different kinds of health-related outcomes can provide information about inconsistent outcomes as potential explanatory triggers. For example, if children anticipate, based on prior knowledge, that individuals will stay healthy, state-changes, such as recovery from illness, may therefore constitute a specific kind of inconsistency with prior experience.

Additionally, like state-change, negative outcomes also often correspond to inconsistent outcomes. For example, children may anticipate or assume (again based on prior knowledge) that outcomes will be positive or favorable, in which case negative outcomes like illness would trigger children's explanations.

In Study 3, preschool children's causal reasoning about illness was investigated, specifically, p*reventing* illness versus *curing* illness. Previous research has shown that children understand that certain factors (such as germs) can intervene to cause illness and other factors (such as medicine) can intervene to cure illness (Kalish, 1996). However, the lack of a change-in-state in the cause-effect relationship involved in the process of illness prevention may make understanding how to maintain health and prevent illness more difficult than reasoning about how to cure or treat illness. In the case of curing illness, the starting-state (illness) is distinct from the end-state (health). However, in the case of illness prevention, the desirable outcome is effectively a null or non-effect, because a person successfully maintains health by avoiding disease. The lack of a perceptible difference or "change-in-state" in the cause-effect relationship involved in the process of illness prevention may influence the conceptual understanding of biological information. Additionally, if children do indeed view maintaining health as consistent with prior knowledge, they may also be less compelled to explain prevention because it lacks an inconsistent outcome.

This distinction is analogous to reasoning about how something generally works (health) versus reasoning about how something goes wrong (illness). In the case of illness, there is a negative outcome, a problem to be solved, and a more general question of something to be addressed or "fixed". However, in the case of preventing illness, it is necessary to reason about effectively sustaining what is essentially a non-effect, in addition to understanding how the body generally functions to maintain health. Indirect evidence that negative or problematic outcomes trigger children's causal explanations comes from research on children's questions about emotions. For example, Wellman and Lagattuta (2004) found that children asked more questions about negative emotions than positive emotions. This suggests that negative events may provoke children's explanations more than positive events or events in which things are functioning as expected. Additionally, there is evidence that counterfactual thinking may be invoked to interpret negative events as part of the causal reasoning process (German, 1999). For these reasons, I anticipate that understanding how to maintain health and prevent illness may be less compelling than reasoning about how to cure or treat illness.

Most developmental research has focused on children's understanding of illness triggers and cures and there is surprisingly little research on children's understanding of illness prevention (Piko & Bak, 2006). The limited research that exists on this topic focuses primarily on children's ability to identify risk factors for disease. For example, research on children's understanding of cancer indicates that although children can identify some risk factors (such as cigarette smoking), more children cite contact or causal contagion than accurate causes such as poor diet, pollution, and sun exposure (Chin, Schonfeld, O'Hare, Mayne, Salovey, Showalter, & Cicchetti, 1998). Yet practically speaking, understanding prevention could have important health benefits.

In order to investigate state-change and negative outcomes as potential causal explanatory triggers, preschool children were presented with contrastive scenarios concerning health and illness and asked to provide explanations. In Studies 1 and 2, the outcome the child explained first was one of the primary measures of children's interest, and a gauge of which outcome was most likely to provoke children's explanations. Therefore, the scenario the child explained first was used as an analogous form of measurement in Study 3. Across a set of conditions, I hypothesized that children would be more likely to first explain a health-state change (becoming sick; curing illness) than a non-state change (remaining sick; preventing illness). I also hypothesized that children would be more likely to first explain a negative outcome (becoming sick; remaining sick) than a positive outcome (curing illness; preventing illness). Finally, if both of these factors (change of state and negative outcome) are relevant, then children should be most compelled to explain a scenario in which a character is initially well and then becomes sick, and least compelled to explain a scenario in which a character is initially well and remains well (prevention).

As in Studies 1 and 2, the content of children's explanations was also of interest. The explanations children provided for health-related outcomes were coded for several kinds of explanatory content, specifically whether children were able to generate relevant explanations for health-related behavior or action. In order to ensure consistency across the vignettes, a sample of adults also participated in Study 3. The expectation was that the adults would perform near ceiling on the task but also that the content of adults' explanations would provide a baseline comparison to children's explanations.

Method

Participants

Participants were 28 preschool children (*M* age 4,3; range 3,1 to 5,2). The sample was recruited from a Midwestern university town, and was primarily White. For comparison purposes a sample of 16 college undergraduates also participated.

Design

A morning-night metric was used across conditions in order for preschoolers to understand a substantial passage of time. For condition 1, a character who was sick in the morning and became healthy at night was contrasted with a character who was healthy in the morning and remained healthy at night. This condition specifically tests the focal contrast between cure and prevention by isolating state-change. I predicted that children will be more compelled to explain the state-change outcome (cure). For condition 2, a character who was sick in the morning and became healthy at night was contrasted with a character who was sick in the morning and remained sick at night. Notably, condition 2 represents an explicit contrast between a state-change and a negative outcome. Therefore I anticipated that because each event represents one of these factors, they may be equally deserving of explanation. For condition 3, a character who was healthy in the morning and became sick at night was contrasted with a character who was healthy in the morning and remained healthy at night. Because the health outcome of becoming sick involves both a state-change and a negative outcome and prevention involves neither, I predicted that children would be much more compelled to first explain the character that became sick. For condition 4, a character who was healthy in the morning and became sick at night was contrasted with a character who was sick in the morning and remained sick at night. As in condition 1, because the valence of both outcomes was the same (in this case both negative), state-change was isolated (Table 6).

Materials and Procedure

Each participant was presented with 2 vignettes in each of 4 different conditions (8 vignettes total). The vignettes were presented in a separate random order for each participant. In each vignette two characters were contrasted, one experienced a health-state change and one did not. Notably, in contrast to Studies 1 and 2, due to the pragmatics of asking participants to explain two distinct health-related events with either dissimilar start-states or dissimilar end-states, it was not possible to present information about both events simultaneously. Presenting the pattern of evidence for each kind of outcome required a substantial amount of descriptive information, as opposed to the simultaneous, visual presentation of both perceptual outcomes in Studies 1 and 2. Therefore, the order the outcomes were presented in (first or last) was counterbalanced.

Each vignette was accompanied by 2 pictures each of 2 different characters. A yellow posterboard representing daytime was placed in front of the participant. Then a picture of the first character was placed on the board followed by a short description of his/her state of health. The pictures corresponded to the character's state of health. For example, in condition 1, a healthy child looked happy and healthy and a sick child looked unhappy with flushed cheeks (see Appendix). Children then heard (for example), "This is Ben in the morning. Ben is sick. He doesn't feel good and his body feels bad". Then a picture of the second character was placed on the board followed by a description of his/her state of health. For example, "This is Jim in the morning. Jim is healthy. He feels good and his body feels strong".

A blue posterboard representing nighttime was then placed on the table below the "day-time" board. The experimenter indicated that a whole day had gone by and that now it was nighttime. To represent the state-change two different pictures of the same characters were used. A picture of the first character was then placed on the "night-time" board below the picture of the character in the morning. A short description of his/her state of health was then provided accompanied by a brief reminder of their prior health-state. For example, "This is Ben at night. Ben was sick in the morning but now he's healthy. Now he feels good and his body feels strong". The same procedure was repeated for the second character. For example, "This is Jim at night. Jim was healthy in the morning and now he's still healthy. He still feels good and his body feels strong". These dual-character presentations were followed by two general questions, "Why is that?" and, "Anything else?"

After the participant provided an explanation, specific questions about each character were asked. This feature of the design was notably different from Studies 1 and 2, in which only ambiguous causal questions were asked, and non-directed explanatory prompts were given. In contrast, in Study 3, after participants were asked an initial ambiguous causal question (referring to either outcome), participants were asked specifically to provide an explanation for each outcome. The motivation for asking directed questions about each outcome was to determine if children would be able to generate a relevant explanation for each kind of health-related outcome (as opposed to whether participants were more likely to find a particular outcome worthy of an explanation at all). For example, a question concerning the character who was sick in the morning and healthy at night was, "Why did that happen? Ben was sick in the morning but now he's healthy, why is that? Why is Ben healthy now?" A question concerning the character who was healthy in the morning and healthy at night was, "Why did that happen? Jim was healthy in the morning and now he's still healthy, why is that? Why is Jim still healthy now?" See the Appendix for sample vignettes for conditions 2-4.

Transcription and Coding

Interviews were videotaped and transcribed verbatim. As in Studies 1 and 2, a primary dependent variable was the outcome the child explained first (state-change =1, non-state-change=0, both=0.5). The other dependent variable was the kind of explanation provided for each kind of health-related outcome. Collapsing across condition, participants had an opportunity to explain each kind of outcome (prevention, cure, remaining sick, and becoming sick) a total of 4 times each. That is, each kind of outcome appeared in half of the conditions and each condition had two vignettes. The content of their explanations for each outcome was coded into five categories: positive action, negative action, absence of positive action, absence of negative action, and unknown. Presence of a specific category of explanation was coded as 1, absence of a specific category of explanation was coded as 0. Explanations were coded as *positive action* if they referred to a specific health-promoting activity (eating fruits and vegetables, getting lots of sleep). Explanations were coded as *negative action* if they referred to a specific health-compromising activity (ate bad food). Explanations were coded as *absence of positive action* if they said that the character failed to engage in a health-promoting activity (not getting enough sleep). Explanations were coded as *absence of negative action* if they said that the character did not engage in a health-compromising activity (not eating bad food). Finally, explanations were coded as *unknown* if participants indicated that they could not provide an explanation or did not know.

Inter-rater reliability was established using a randomly selected sample of 25% of the data. Two persons independently coded the reliability sample with 94% agreement. Reliability was calculated for explanatory coding categories for positive action, negative action, absence-of positive action, absence-of-negative action, and don't know explanatory categories; with Kappas ranging from .93 to .81. All of the Kappas for this coding fall within near perfect (.80 and above) levels (Landis & Koch, 1977).

Composite coding categories

In order to investigate whether participants were more able to provide relevant explanations for some health-related outcomes than others, summary score variables were computed based on whether the kind of explanation provided for each kind outcome was a "match", "mismatch", or "unknown". Explanations were coded as a "match" if the kind of explanation given was appropriate and relevant for the outcome. For example, both a negative action and a lack of a positive action were coded as a "match" explanation for a character who became sick or remained sick. Explanations were coded as a "mismatch" if the kind of explanation given was inappropriate for the outcome. For example, both a positive action and a lack of a negative action were coded as a "mismatch" explanation for a character who became sick or remained sick (Table 7). Finally, "unknown" explanations were included in an additional category for all four kinds of state-change and non-state-change outcomes. Children had an opportunity to provide an explanation for each kind of health-related outcome 4 times (collapsing across conditions) across a total of 8 vignettes.

Comparing the content of participants' explanations for state-change and negative outcomes

In order to determine if participants were more able to generate relevant explanations for state-change and negative outcomes than non-state-change and positive outcomes, summary score variables were computed. These variables were calculated by adding together "match" explanation scores for each kind of outcome (state-change outcomes, non-state-change outcomes, negative outcomes, and positive outcomes). For example, "match" explanation scores for becoming sick were added to scores for cure to calculate a state-change score. Explanation scores for remaining sick were added to scores for prevention to calculate a non-state-change score. Explanation scores for prevention and cure were added together to calculate positive outcome scores. Explanation scores for becoming sick and remaining sick were added together to calculate negative outcome scores. "Mismatch" scores and "unknown" summary scores were calculated the same way. Participants had an opportunity to provide an explanation for each kind of health-related outcome 4 times (each outcome appeared in two conditions and each condition had two vignettes that counterbalanced the order the outcome was presented in). By adding two kinds of outcomes together, scores were out of a total of 8 possible (Table 8).

Results

Participants' explanations for health-state changes were contrasted with non-state changes in order to determine whether participants were more likely to explain health-state changes first (Table 6). I hypothesized that although state-changes would be explained first most often, other explanatory triggers, such as negative outcomes would also provoke children's explanations. Conditions 1-4 were designed to test a priori planned comparisons with distinct predictions that varied by condition, therefore t-test comparisons to chance were used to calculate which health-related outcome participants explained first, instead of an omnibus ANOVA test. Data from the preschool children will be presented first, followed by the adult participants. The results of the content of participants' explanations will follow the results of which outcome participants explained first.

What did preschool children explain first?

In all conditions in which a state-change was contrasted with a non-state-change outcome and the valence of the outcomes was held constant (both positive outcomes (condition 1) and both negative outcomes (condition 4)), children were more likely to explain the state-change outcome first. Additionally, when a state-change outcome that coincided with a negative outcome (became sick) was contrasted with a positive, non-state-change outcome (prevention) (condition 3), children were much more likely to explain the negative, state-change outcome first. In contrast, when a state-change outcome (cure) competed with a negative outcome (remain sick), first explanation was at chance. Although the means for first explanation in conditions 1-4 were not significantly different from each other, the means are in the right direction. That is, the mean for condition 3 is highest, followed by conditions 1 and 4, and finally condition 2 (Figure 9).

State-change versus non-state-change outcomes

The focal contrast between cure and prevention isolated state-change (the valence of both outcomes was positive) (Table 6). I predicted that children would be more compelled to first explain a cure than illness prevention. As hypothesized, in condition 1, when simply asked, "why did that happen" without specifically referring to one character or another, children were more likely to first explain a character who was sick in the morning and became healthy at night (cure) than a character who was healthy in the morning and remained healthy at night (prevention), $t(27) = 2.5$, $p < .05$ (Figure 9). Note that if children were merely attending to health outcome, children's first explanation for these contrastive events should have been at chance given that cure and prevention outcomes were equivalent (i.e., in both cases the character ended up healthy).

When the valence of both outcomes was negative and state-change was isolated (condition 4) (Table 6), children were more likely to explain the state-change outcome first. That is, children were more likely to first explain a character who was healthy in the morning and became sick at night (became sick) than a character who was sick in the morning and remained sick at night (remain sick), $t(27) = 2.1$, $p < .05$. Based on the equivalent health outcome (illness), the greater likelihood of explaining a health state-change first provides further evidence that children are more compelled to explain state-changes than non-state-changes.

Additionally, as predicted, when a positive, non-state-change outcome (prevention) was contrasted with a negative, state-change outcome (became sick), children were also much more likely to explain the negative, state-change outcome first, $t(27) = 3.3$, $p < .01$(condition 3). If children were merely attending to the initial health state, children's first explanation for these contrastive events should have been at chance given that both characters started out healthy. This also provides further evidence that explaining a health state-change is more compelling to children than illness prevention (a non-state-change).

<u>State-change versus negative outcome</u>

I anticipated that when state-change was explicitly contrasted with a negative outcome (condition 2), both outcomes would be equally likely to provoke children's explanations and therefore that the outcome they explained first would be at chance. The results confirm this prediction. Children were equally likely to first explain a character who was sick in the morning and became healthy at night (cure) and a character who was sick in the morning and remained sick at night (remain sick), $t(27) = 1.14$, *ns* (Figure 9). This suggests that, as anticipated, although state-change is an important factor in what children feel compelled to explain, other factors, such as negative outcomes (illness), are also important.

What did adults explain first?

A sample of adults participated in Study 3 in order to ensure consistency across the vignettes. Although the expectation was that the adults would perform near ceiling on the task, I was also interested in whether the adults' data would show the same bias to first explain a state-change or negative health-related outcome that I found in preschool children. However, the data indicate that adults were equally likely to first explain outcomes involving a state-change as a non-state-change in this task. Across conditions, there were no significant differences in which kind of outcome they explained first. One possibility for this result is that adults may be more likely than children to consider both state-change and negative outcomes when providing explanations for health-related outcomes. Another possibility is that this task may not be sensitive enough to detect these potential explanatory biases in adults.

The content of explanations for different kinds of health-related outcomes

In addition to investigating whether participants were more likely to first explain state-change or negative outcomes in contrast to non-state-change or positive outcomes, I was also interested in investigating whether participants were more able to provide relevant explanations for health-related outcomes involving a state-change or negative valence. The prediction was that participants would provide more relevant explanations for health-state changes than non-state changes. Therefore, the content of the explanations participants provided for each kind of health-related outcome was analyzed. In order to do so, the number of explanations that fell into each of the five explanatory categories was calculated. Collapsing across condition, summary scores for both state-change versus non-state-change and negative versus positive outcomes were computed. To examine potential effects of state-change outcomes on participants' explanations, explanation types for each health-state change (becoming sick and being cured) and non-state change (prevention and remaining sick) were compared using paired t-tests. Additionally, to examine potential effects of negative versus positive outcomes on participants' explanations, explanation types for each negative outcome (becoming sick and staying sick) and positive outcome (prevention and cure) were analyzed in the same way.

Content of preschool children's explanations

<u>State-change versus non-state-change outcomes</u>

Preschool children provided more relevant explanations for health-state changes than non-state changes. That is, children were significantly more likely to give "match" explanations to explain state-change outcomes (becoming sick and cure) (M=5.96 out of 8 possible) than non-state-change outcomes (staying sick and prevention) (M=4.39), $t(27) = 3.62$, $p < .001$, indicating that, as predicted, they were more able to generate appropriate or accurate explanations for a state-change outcome than a non-state-change outcome (Figure 10). Nonetheless, children gave significantly more "match" explanations than "mismatch" explanations for both state-change outcomes, $t(27) = 10.43$, $p < .001$, and non-state-change outcomes, $t(27) = 7.69$, $p < .001$, indicating that even young children have access to information about illness and can provide relevant explanations for health-related outcomes. Additionally, the data indicate that "mismatch" explanations were very infrequent overall (M=0.3 out of 8 possible for state-change outcomes and 0.0 for non-state-change outcomes).

In contrast, children were much more likely to fail to generate an explanation (by indicating that they "don't know") for a non-state-change outcome (M=3.61) than a state-change outcome (M=1.71), $t(27) = 4.70$, $p < .001$ (Figure 10). These results support my hypothesis that although young children can generate relevant biological explanations, outcomes that involve a state-change may be more compelling, and conversely, that outcomes lacking a state-change may be less compelling from the standpoint of explanation.

<u>Negative versus positive outcomes</u>

Unlike the results of the analyses of state-change versus non-state-change outcomes, children were equally likely to provide relevant explanations for negative health outcomes as for positive health outcomes. That is, participants were equally likely to give "match" explanations to explain negative outcomes (becoming sick and remaining sick) (M=5.32 out of 8 possible) as positive outcomes (staying healthy and prevention) (M=5.04), $t(27) = .78$, *ns* (Figure 10). However, as in the case of explanations for state-change versus non-state-change outcomes, children gave significantly more "match" explanations than "mismatch" explanations for both negative outcomes, $t(27) = 9.55$, $p < .001$, and positive outcomes, $t(27) = 9.31$, $p < .001$, providing further evidence that young children can provide relevant explanations for health-related outcomes.

Children were also equally likely to fail to generate an explanation (by indicating that they "don't know") for a negative outcome (M=2.46) as a positive outcome (M=2.86), $t(27) = 1.05$, *ns* (Figure 10). These results provide evidence that, relative to state-change, negative outcomes may play less of a role in provoking children's explanations.

Content of adults' explanations

A sample of adults participated in Study 3 in order to establish a baseline comparison to the content children's explanations. Although adults were anticipated to perform at ceiling (and did so), the content of their explanations was highly similar to the content of children's explanations for the same outcomes. This validates the suitability of the coding categories and the content of children's explanations as being insightful and appropriate.

<u>State-change versus non-state-change outcomes</u>

Unlike the child data, adults were equally likely to give “match” explanations for state-change (M=7.81), and non-state-change outcomes (*M*=7.68), $t(15) = 1.00$, *ns,* indicating that, unlike the child sample, adults were equally likely to provide relevant explanations for health-state changes and non-state changes. Failure to generate an explanation for an outcome was very infrequent (M=.19 out of 8 for state-change outcomes and M=.25 for non-state-change outcomes) and equivalent across outcomes, $t(15) = .57$, *ns.* They also gave significantly more “match” explanations than “mismatch” explanations for both state-change outcomes, $t(15) = 57.5$, $p < .001$, and non-state-change outcomes, $t(15) = 49.3$, $p < .001$.

Negative versus positive outcomes

The results of adults’ explanations for negative versus positive health-related outcomes mirror the results of explanations for state-change versus non-state-change. Adults were equally likely to give “match” explanations for negative outcomes (*M*=7.69), as positive outcomes (*M*=7.81), $t(15) = 1.00$, *ns,* indicating that, as with the child sample, adults were equally likely to provide relevant explanations for negative and positive outcomes. Additionally, explanations for “match” explanations were significantly more common than “mismatch” explanations for both negative outcomes, $t(15) = 51.07$, $p < .001$, and positive outcomes, $t(15) = 53.69$, $p < .001$. Failure to generate an explanation for an outcome was very infrequent (M=.31 out of 8 for negative outcomes and M=.13 for positive outcomes) and equivalent across outcomes, $t(15) = 1.86$, *ns.*

Discussion

The data from Study 3 provide evidence that state-change is an especially powerful trigger for children's causal reasoning. As hypothesized, children were more likely to first explain and provide relevant explanations for state-change than non-state-change outcomes. An important example of this is that children were less likely to explain prevention first, and less able to generate an explanation for preventing illness than recovering from illness (cures) and becoming sick. This may be because preventing illness does not involve a state-change and may also be an unsurprising, expected outcome, therefore children may be less compelled to attend to and reason about it.

In three of the four conditions, children were significantly more likely to explain a health-state change first. Notably, in both cases when prevention (a non health-state change) was contrasted with a health-state change (cure or becoming sick), preschool children were more likely to explain the health-state change. This would not be anticipated if children were merely attending to initial health-state, because in the contrast between prevention and becoming sick they are equivalent. Relatedly, attending to health-state outcome also fails to explain this pattern because in the contrast between curing illness and prevention, health state outcomes are equivalent.

Further support for the special role of state-change in provoking causal reasoning was found in the content of the explanations children generated to explain these paired events. Preschool children provided more relevant explanations for health-state changes than non-state changes. Additionally, children were less able to generate an explanation for non-health-state changes by indicating that they didn't know. Importantly, the fact that preschool children were less able to generate relevant explanations for non-health-state changes cannot be attributed to a general lack of ability to explain health or illness. For example, preschool children proved to be very capable of providing appropriate explanations for cures or recovery from illness. Given this, the fact that children were more likely to provide relevant explanations for characters that became sick or became healthy (cured) than characters that remained healthy (prevention) is especially important. Again, this would not be anticipated if participants were simply attending to the initial health-state or outcome health-state because in each comparison case, it was equivalent.

Additionally, although the adult data did not demonstrate the same bias to first explain a state-change or negative outcome as the child data, the content of the adults' explanations was highly similar to the content of children's explanations for the same outcomes. This validates the suitability of the coding categories and the content of children's explanations as being insightful and appropriate.

The data also indicate that children attend to more than state-change when providing explanations for outcomes. Although children were more likely to first explain a character who became sick than a character who remained sick, they were equally likely to first explain a character who was cured of illness as a character who remained sick. One possibility for this result is that children are also attending to negative outcomes and inconsistency with prior knowledge. For example, although a cure involves a noteworthy state-change, remaining sick is both a highly negative outcome and may violate children's expectations for health outcomes. Alternately, in the case of becoming sick versus remaining sick, because both outcomes are negative, children may attend more to the health-state change (becoming sick). However, unlike the results of the analyses of state-change versus non-state-change outcomes, participants were equally likely to provide relevant explanations for negative health outcomes as for positive health outcomes. These results provide evidence that, relative to state-change, negative outcomes may play less of a role in provoking children's explanations.

Although there is a well-established literature on young children's biological reasoning, there is surprisingly little research on children's understanding of illness prevention (Piko & Bak, 2006). Yet practically speaking, understanding prevention could have important health benefits. One potential educational implication of this research is that the process of avoiding or preventing illness may require explicit instruction. My data are also consistent with the possibility that the limited success health education programs have in influencing illness prevention behavior may be due in part to the fact that even at a young age, prevention may be less compelling from a cognitive perspective due to greater interest and attention to outcomes that involve state-change.

Finally, although state-change is a compelling trigger for children's causal explanations, it does not operate in isolation. Negative outcomes may also play a role in the kinds of events children attend to and feel compelled to explain. In future studies, experimentally disambiguating state change from negative outcomes could be targeted more explicitly by using a task similar to the task involving novel causal properties and categories used in Studies 1 and 2.

Chapter V

General Discussion

Causal explanation is a goal-directed cognitive activity. It depends on what is relevant or important to the explainer. Given the potentially limitless number of outcomes and events one could seek to explain, especially for young children early in the process of developing causal knowledge, identifying the specific kinds of events, outcomes, and goals that trigger explanation provides insight into what motivates (and guides) causal cognition. Although a substantial amount of developmental literature has demonstrated that young children frequently seek out (Callanan & Oakes, 1992) and provide explanations (Wellman, Hickling, & Schult, 1997), less is known about what triggers or motivates childhood explanatory activity.

The primary objective of this dissertation was to explore explanatory biases or triggers in preschool children, using their explanations as the primary dependent measure. I propose that explanatory biases play an important role in guiding children's causal explanations. I investigated three interrelated, but conceptually distinct kinds of outcomes that may potentially trigger causal explanation in children: outcomes inconsistent with prior knowledge, state-change outcomes, and negative outcomes. Because it is possible that all three kinds of outcomes may provoke causal explanation in children and because these kinds of outcomes are often confounded both in the world and in experimental manipulations, a primary motivation behind the design of Studies 1-3 was to experimentally differentiate the kinds of events that children find noteworthy and therefore feel compelled to explain.

If children have a cognitive model of the world based on a framework of anticipatory causal regularities, they would be well-equipped to rapidly form expectations contingent upon prior beliefs or knowledge, even from very limited available input. Because explanations are dependent to some extent on prior knowledge (Medin, Coley, & Storms, 2003; Sloman, 1994), constructing causal explanations is a mechanism through which prior knowledge serves as a basis for interpreting a new outcome. Children construct explanations for events that are *consistent with* prior knowledge but also recognize and attempt to explain events that are *inconsistent* with prior knowledge. For this reason, children's explanations may serve at least two distinct functions. If explanation is largely confirmatory, children should be motivated to construct explanations for outcomes that are consistent with prior knowledge. That is, if children have a predisposition to prognosticate causal regularities, children may anticipate that outcomes will continue to occur as expected and find consistent outcomes especially worthy of explanation. If, on the other hand, explanation is a mechanism for discovery, children should be motivated to construct explanations for events that are inconsistent with their prior knowledge.

Thus, the primary objective of Studies 1 and 2 was to experimentally investigate these two competing hypotheses about the function of children's explanations by determining whether children reason differently about consistent events versus events in which something inconsistent happens. Data confirm my prediction that events inconsistent with prior knowledge trigger children's causal explanations. Results from Studies 1 and 2 demonstrate that outcomes inconsistent with prior beliefs both attract children's attention and provoke children's explanations. Children were much more likely to explain inconsistent outcomes than consistent outcomes, even though they (a) had multiple opportunities to explain both, and (b) had access to an a priori explanation for the consistent outcome. Their explanations also went beyond surface features to include information about causal mechanisms and redefining category membership, even when object labels were novel and did not provide additional functional information. These data provide evidence that children are not simply using explanation to confirm what they already know, but instead provide evidence that children use explanation in the process of discovering new information and constructing new understanding.

An additional objective of Studies 1 and 2 was to investigate the possibility that what is relevant to children's explanations is neither consistency nor inconsistency, but other factors (such as state-change or negative outcomes). An important feature of the design of the test trials in Studies 1 and 2 was to both control for additional potential explanatory triggers (state-change and negative outcomes) by holding them constant, and to investigate state-change as a potential explanatory trigger in the absence of an inconsistent outcome (confirmation trials). Results indicate that although inconsistency with prior knowledge is a compelling trigger for children's causal explanations, in the absence of this information, state-change is an additional explanatory trigger.

Study 3 expanded on Studies 1 and 2 by focusing on state-change and negative outcomes as potential causal explanatory triggers in the biological domain, specifically concerning health-related outcomes. Study 3 focused specifically on children's explanations of health-related outcomes because state-change and negative valence are both meaningful and relevant in the domain of illness. Data from Study 3 indicate that as hypothesized, children are more likely to first explain and provide relevant explanations for state-change outcomes than non-state-change outcomes. Additionally, the data support my prediction that because preventing illness does not involve a state-change, children may be less compelled to attend to and reason about it. For example, children were less able to generate explanations for preventing illness than recovering from illness (cures) and becoming sick.

Although the idea that inconsistent, problematic, or surprising outcomes play an important role in causal reasoning appears across multiple literatures, including infant cognition research (Baillargeon, 2002), there is remarkably little empirical research on what motivates causal explanatory reasoning in children and how this can inform the developmental trajectory of causal explanation. The objective of examining explanatory triggers with preschool children is not simply to use their verbal articulation to demonstrate that they attend to inconsistent or problematic outcomes. Rather, the objective is to investigate the possibility that explanatory triggers go beyond orienting children's attention by provoking children to construct causal explanations and new kinds of causal knowledge. My data provide strong evidence that both inconsistent outcomes and state-change outcomes motivate preschool children to generate causal explanations, thus supporting and extending prior work done with infants.

Educational Implications

There is mounting evidence that children's causal explanations may in fact constitute a mechanism for advancing causal learning and the acquisition of knowledge (Chi et al., 1989). Indeed, other studies suggest that requiring children to explain events enhances learning over simple feedback about correctness of their predictions (e.g., Amsterlaw & Wellman, 2006; Siegler, 1995). Thus, explanations not only reveal causal reasoning, they seem to play an important role in helping children discover and learn new ways of thinking (Chi, DeLeeuw, Chiu, & LaVancher, 1994; Siegler, 1995). But how so? To the extent that explanations may represent children's most advanced theoretical reasoning, they may provide an important platform for further understanding and learning. My data are consistent with the possibility that children's explanations may play an active, important role in the learning process. If explanation is a mechanism for learning, children should provide explanations for events that have the potential to teach them something new. Events that are inconsistent with prior knowledge provide just such an opportunity. Explanatory tendencies may aid in learning by providing a mechanism for advancing causal knowledge acquisition.

The nature of the explanations I observed, in appealing to unobserved explanatory processes and functional properties, could also be important for learning of the sort described by a naïve theory perspective on cognitive development (Gelman & Wellman, 1991). To the extent that explanations appeal to underlying, internal causal properties and mechanisms, they engage children in the important interplay between data and theory that leads to theory change. Although these particular studies were not designed to investigate explanation-based mechanisms for causal learning, given my results, this becomes an important topic for further research.

Additionally, a potential implication of my research is that the kinds of events that motivate causal explanation may have important educational consequences. For example, given the importance of prevention in health education programs, my data suggest that the process of avoiding or preventing illness requires special emphasis and explicit instruction. My results are also consistent with the possibility that the limited success health education programs have in influencing illness prevention behavior may be due, in part, to the fact that even at a young age, illness prevention may be less compelling from a cognitive perspective than health-related outcomes that involve a state-change or negative outcome.

Limitations and directions for future research

The objective of this dissertation was to begin to experimentally disentangle the kinds of events that trigger children's causal explanations. In Studies 1 and 2, inconsistency with prior knowledge was of primary interest and therefore state-change and the valence of the outcome was held constant. However, in future studies, state-change and negative outcomes could be specifically targeted by holding inconsistency with prior knowledge constant using the same task involving novel causal properties and categories used in Studies 1 and 2.

Additionally, data from Study 3 indicate that children may have been attending to more than state-change when providing explanations for outcomes. For example, children were equally like to provide an explanation for a character who was cured of illness as a character who remained sick. One possibility for this result is that children are also attending to negative outcomes and inconsistency with prior knowledge. Whereas a cure involves a noteworthy state-change, remaining sick is a highly negative outcome and may violate children's expectations for health-based outcomes. Alternately, in the case of becoming sick versus remaining sick, because both outcomes are negative, children may attend more to the health-state change (becoming sick). Although limitations in the design of Study 3 did not allow for specifically targeting negative outcomes as a potential explanatory trigger, this would be an interesting direction for future research.

One of the strengths of Studies 1 and 2 was that the design allowed for the simultaneous presentation of contrastive outcomes. In Study 3, due to the pragmatics of asking participants to explain two distinct health-related events with either dissimilar start-states or dissimilar end-states, it was not possible to present information about both events simultaneously. Although this was a limitation of the design of Study 3, it was accommodated by counterbalancing the order the outcomes were presented in (first or last).

The relationship between children's explanations and their exploratory play is another interesting direction for future research. Recent work by Schulz & Bonawitz (2007) shows that children engage in more exploratory play when evidence is confounded than when it is not, and Bonawitz, Lim, & Schulz (2007) demonstrate that children are more likely to explore an object that functions in a way that conflicts with prior beliefs rather than exploring a novel object. Additionally, data from Bonawitz et al. (2007) indicate that children generate different explanations when evidence conflicts with their prior beliefs than when it confirms their prior beliefs. In ongoing studies I am extending upon my dissertation research by investigating whether the kinds of explanations children provide influence their exploratory play, and how children use explanation in the process of hypothesis-testing and problem solving.

In sum, data from three studies provide evidence for the interplay of three distinct, but interrelated biases that guide children's causal explanatory reasoning. The data also provide insight into the function of children's explanations and provide empirical evidence for the kinds of events that guide causal cognition by motivating children to construct explanations.

Table 1: Studies 1 and 2: Study Design:

Condition	**Generative (Off→X)**						**Inhibitory (On→X)**					
Trial	CT		TT		TT		CT		TT		TT	
Initial State of Light Box	Off	Off	Off	Off	Off	Off	On	On	On	On	On	On
Type of Object	Starter	Do-nothing	Starter	Do-nothing	Starter	Do-nothing	Stopper	Do-nothing	Stopper	Do-nothing	Stopper	Do-nothing
Final State of Light Box	On	Off	Off	Off	On	On	Off	On	Off	Off	On	On
Inconsistent Outcome=1	0	0	1	0	0	1	0	0	0	1	1	0
State-change=1	1	0	0	0	1	1	1	0	1	1	0	0
Negative Outcome =1(off)	0	1	1	1	0	0	1	0	1	1	0	0
Isolates	State-change/ Negative Outcome		Inconsistent Outcome		Inconsistent Outcome		State-change/ Negative Outcome		Inconsistent Outcome		Inconsistent Outcome	

Table 2: Study 1: Sample Causal Explanations for Inconsistent Outcomes

Category switch	**Causal function**	**Causal action**
"It looks like there was one do-nothing, and it seems like it wasn't going to start but it did start. It's really a starter"	"The stopper isn't working anymore, it's broken inside"	"You set the starter on the wrong side"
"It looks like that one was a do-nothing, but it wasn't a do-nothing it was a stopper. It makes the box stop"	"Because this went out of power. Because it doesn't have stoppers in it"	"The stopper should be on the other box, that's why it isn't working right"
"Because that is a starter too. Because some are starters too. Some of these are starters too"	"Because it's broken. I think because it doesn't have batteries in it"	"They are on the wrong boxes"

Table 3: Study 2: Sample Causal Explanations for Inconsistent Outcomes

Causal category	**Causal function**	**Causal action**
“It’s a toma, it just looks like it isn’t one.”	“The toma isn’t working anymore, it’s broken inside”	“You set the blicket on the wrong side”
“It turned on because it is a blick”	“Maybe because this one has more air and energy”	“The toma should be on the other box, maybe try putting it on there”
“Because that one can make it turn off. It is the same kind of thing as the toma”	“The toma didn’t stop. It didn’t stop because it doesn’t have any magnets in it.”	Because that one (blicket) is supposed to be on the other side and the one that isn’t a blicket is supposed to be on the other box”

Table 4a: Study 1: Percentage of Children Giving Each Kind of Causal Explanation for Inconsistent Outcomes at Least Once by Age Group

	Category switch	Causal function	Causal action
3-year-olds	31%	56%	19%
4-year-olds	56%	69%	6%
5-year-olds	69%	56%	19%

Table 4b: Study 2: Percentage of Children Giving Each Kind of Causal Explanation for Inconsistent Outcomes at Least Once by Age Group

	Category switch	Causal function	Causal action
3-year-olds	25%	94%	13%
4-year-olds	69%	50%	19%

Table 5a: Studies 1 and 2: Percentage of Children Providing More than One Kind of Causal Explanation for Inconsistent Outcomes by Age Group

	Study 1	Study 2
3-year-olds	19%	31%
4-year-olds	25%	38%
5-year-olds	50%	-------

Table 5b: Studies 1 and 2: Percentage of Children Providing More than One Kind of Causal Explanation for Consistent Outcomes by Age Group

	Study 1	Study 2
3-year-olds	19%	13%
4-year-olds	38%	25%
5-year-olds	50%	-------

Table 6: Study 3: Study Design

		State-change	
Non-state-change		**Positive outcome** Sick-Healthy (cure)	**Negative outcome** Healthy-Sick (become sick)
	Positive outcome Healthy-Healthy (prevention)	Condition 1	Condition 3
	Negative outcome Sick-Sick (remain sick)	Condition 2	Condition 4

Table 7: Study 3: Coding Categories for the Kinds of Explanations Provided for Health-related Outcomes

	Match	**Mismatch**
State-change: Healthy-Sick (become sick)	Negative action, lack of positive action	Positive action, lack of negative action
State-change: Sick-Healthy (cure)	Positive action, lack of negative action	Negative action, lack of positive action
Non-state-change: Healthy-Healthy (prevention)	Positive action, lack of negative action	Negative action, lack of positive action
Non-state-change: Sick-Sick (remain sick)	Negative action, lack of positive action	Positive action, lack of negative action

Table 8: Study 3: Summary Score Variables for State-change and Negative Outcomes

Becoming sick + Cure	Change-of-state
Remaining sick + Prevention	Non-state-change
Cure +Prevention	Positive outcome
Becoming sick + Remaining sick	Negative outcome

Figure 1: Studies 1 and 2: Object Stimuli

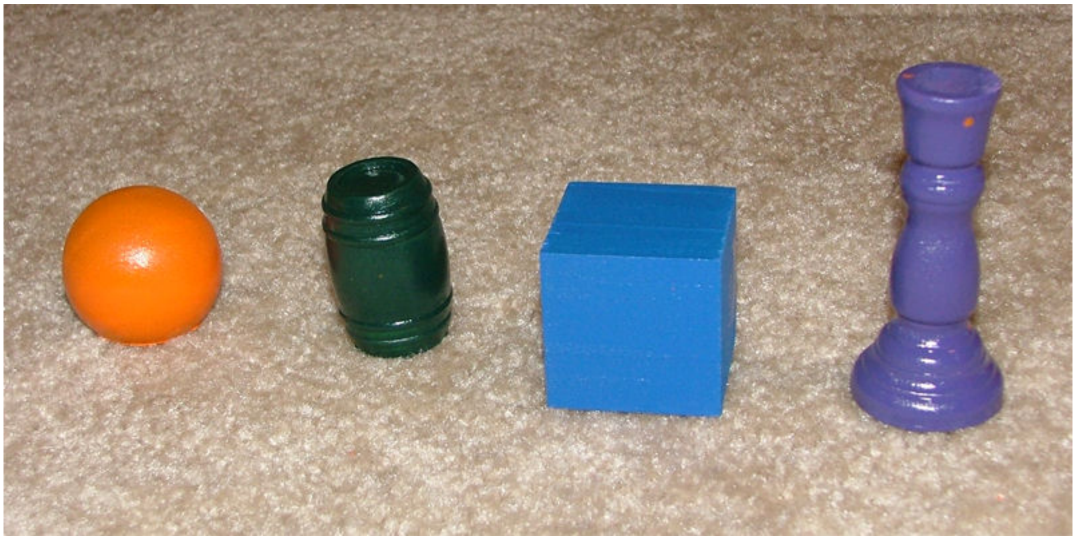


Figure 2: Studies 1 and 2: Generative Condition

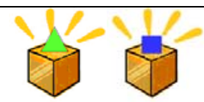
Generative condition	 = starter/toma  = do-nothing/not-a-toma
Confirmation trial	
Test trial: Light boxes both turn on	
Test trial: Light boxes both stay off	

Figure 3: Studies 1 and 2: Inhibitory Condition

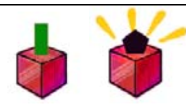
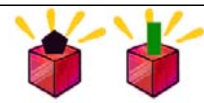
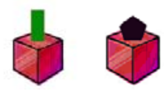
Inhibitory condition	 = stopper/blicket  = do-nothing/not-a-blicket
Confirmation trial	
Test trial: Light boxes both stay on	
Test trial: Light boxes both turn off	

Figure 4: Studies 1 and 2: Two Distinct Hypotheses about the Function of Causal Explanation

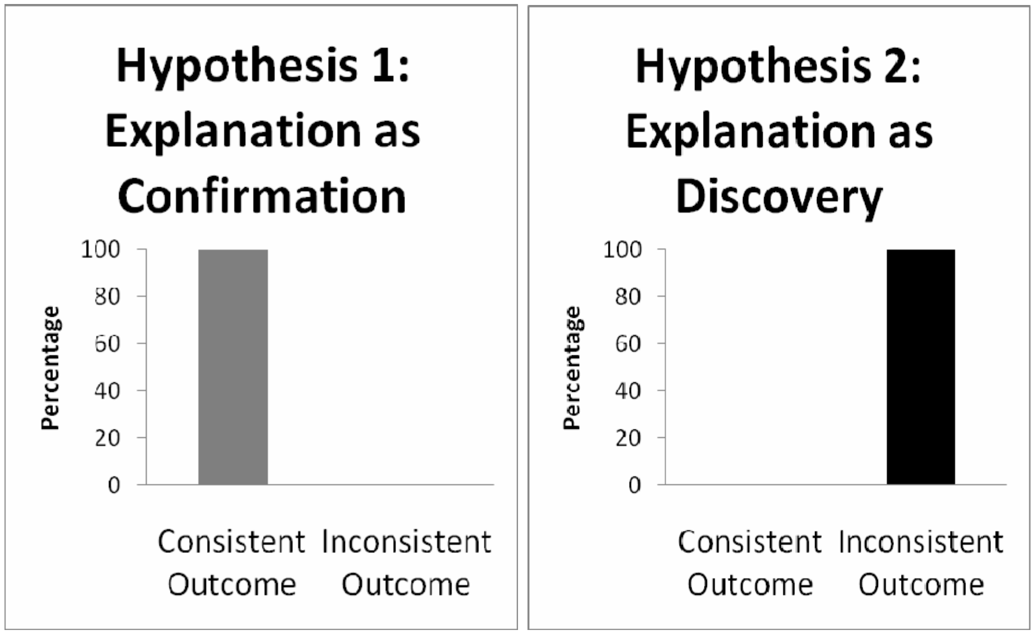


Figure 5: Studies 1 and 2: First Explanation

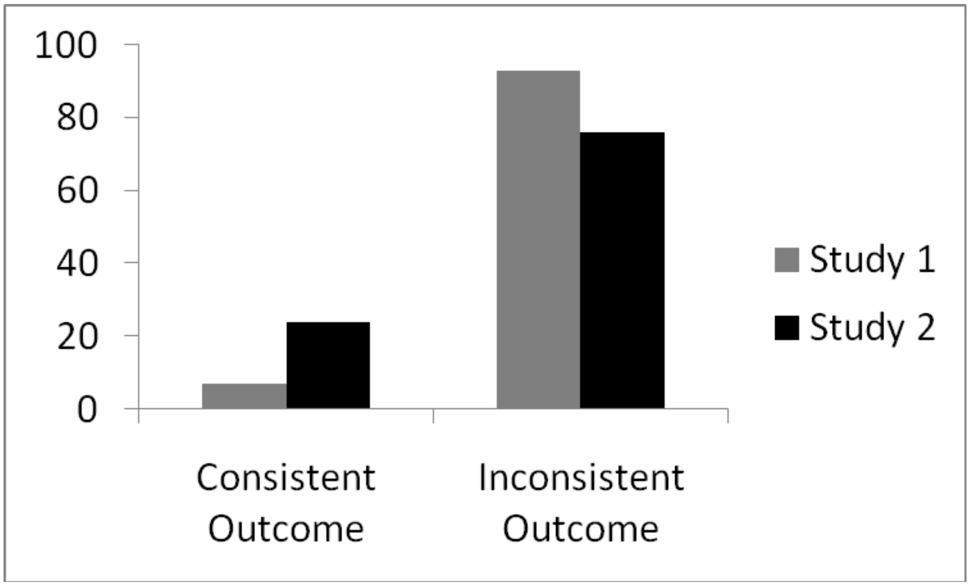


Figure 6: Studies 1 and 2: All Explanations

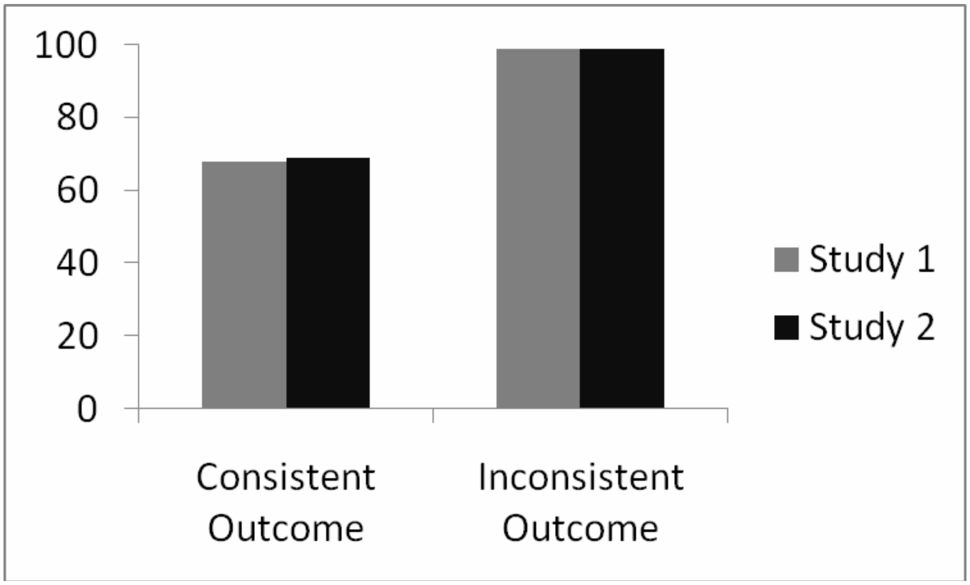
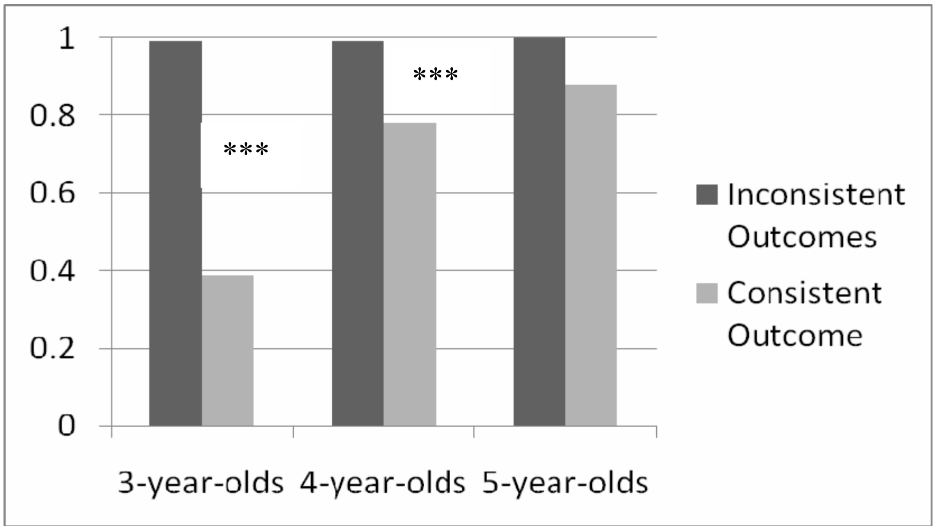
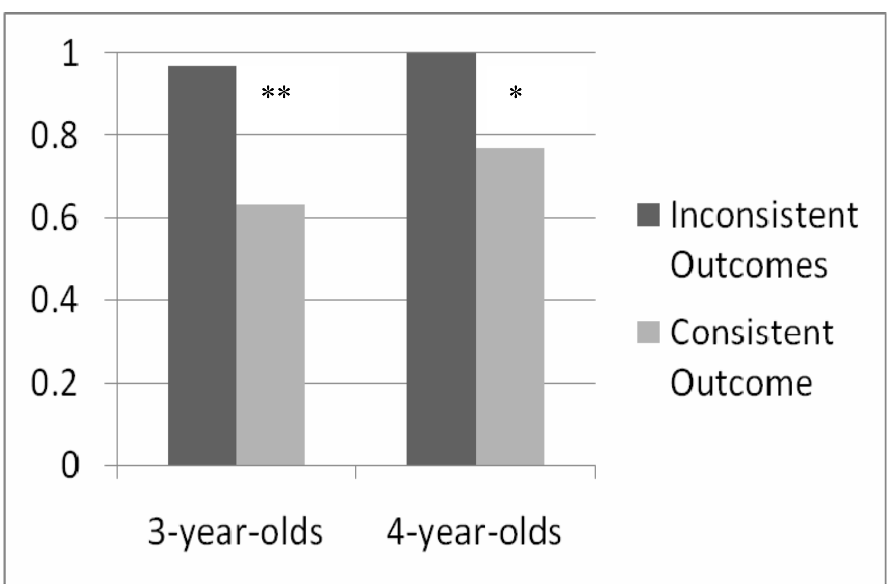


Figure 7a: Study 1: Proportion of Explanations Children Provided for Inconsistent and Consistent Outcomes by Age Group



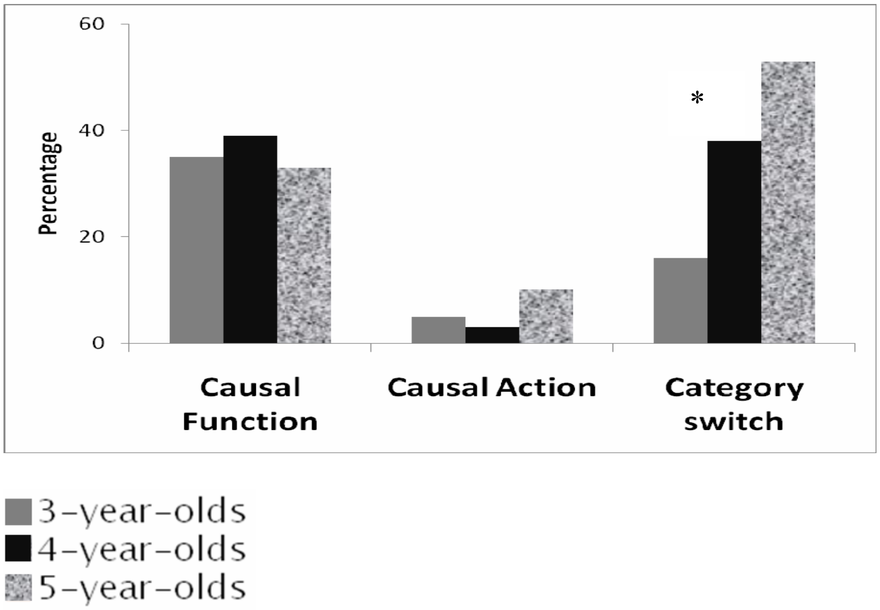
*** *p*>.001

Figure 7b: Study 2: Proportion of Explanations Children Provided for Inconsistent and Consistent Outcomes by Age Group



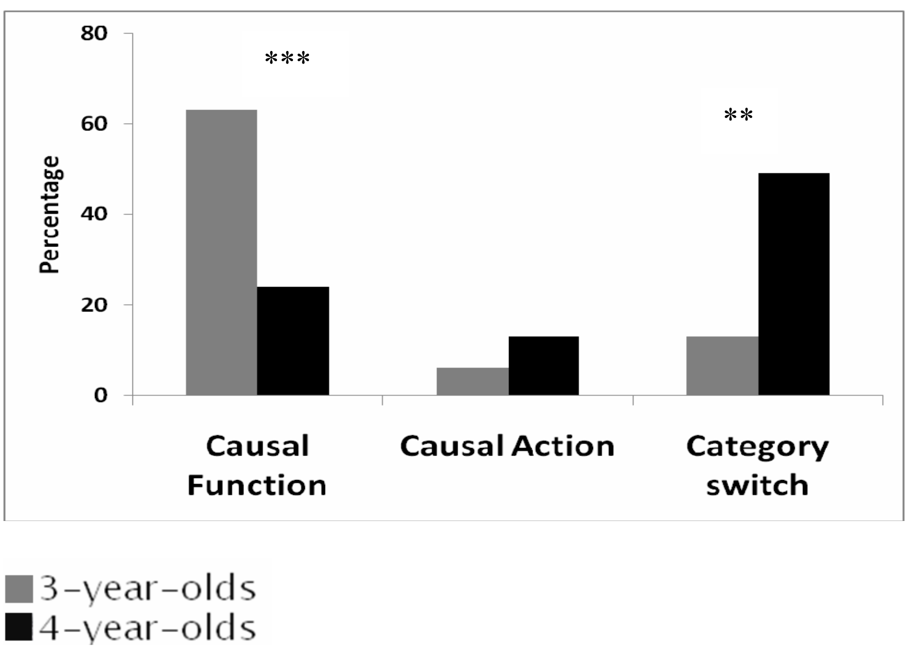
* *p*>.05, ** *p*>.01

Figure 8a: Study 1: Percentage of Total Explanations for Inconsistent Outcomes by Age



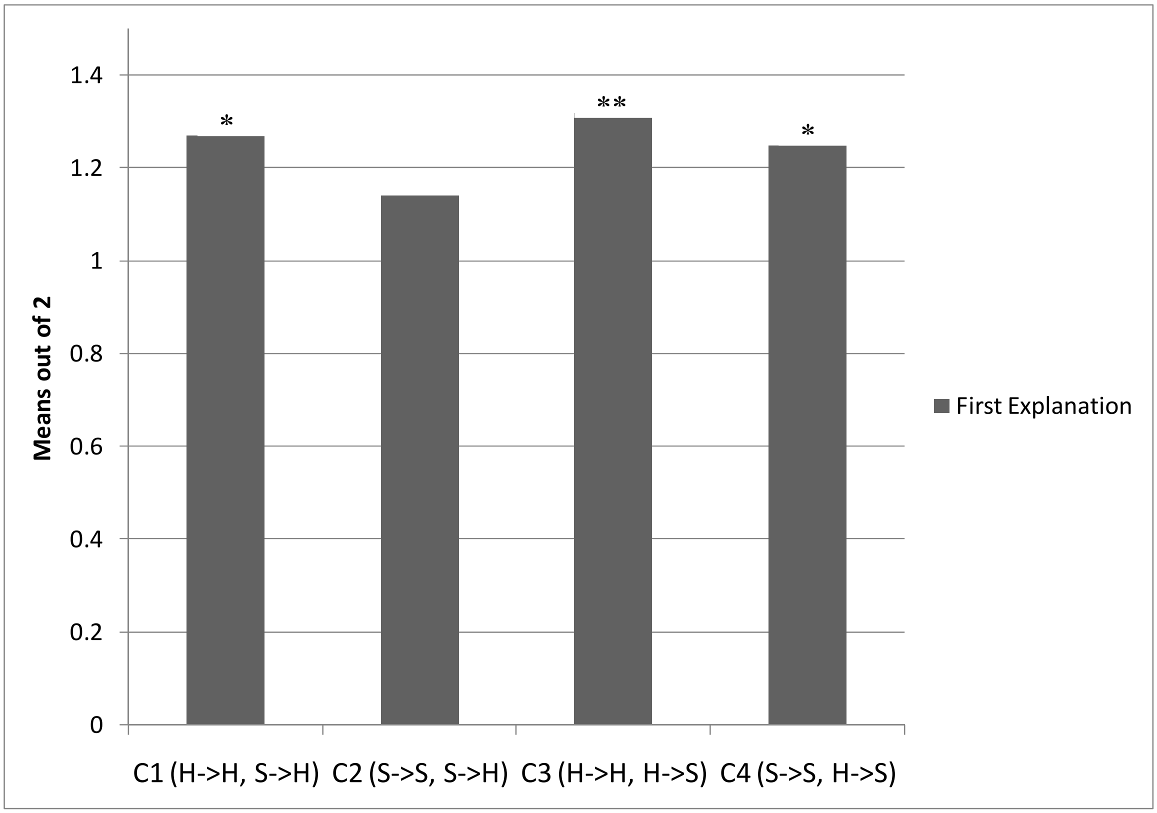
* *p*>.05

Figure 8b: Study 2: Percentage of Total Explanations for Inconsistent Outcomes by Age



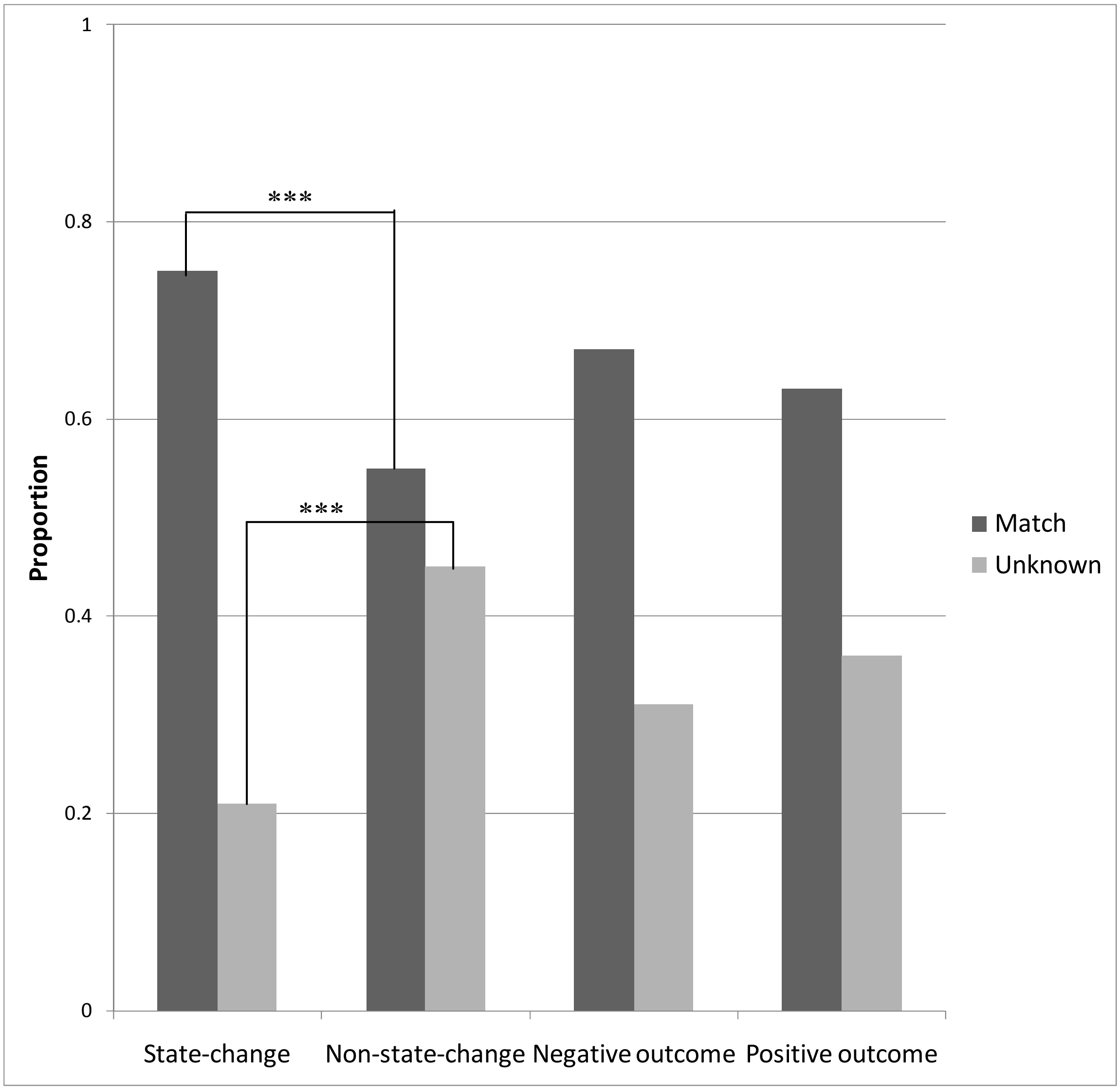
** *p*>.01, *** *p*>.001

Figure 9: Study 3: First Explanation: State-Change



* *p*>.05, ** *p*>.01

Figure 10: Study 3: Proportion of Total Explanations for Health-related Outcomes



*** *p*>.001

Appendix

Study 3: Sample Vignettes

Sick to healthy, healthy to healthy:

This is Ben in the morning. Ben is sick. He doesn't feel good and his body feels bad. This is Jim in the morning. Jim is healthy. He feels good and his body feels strong. A whole day goes by. Now it's nighttime. This is Ben at night. Ben was sick in the morning but now he's healthy. Now he feels good and his body feels strong. This is Jim at night. Jim was healthy in the morning and now he's still healthy. He still feels good and his body feels strong.

- Why is that? Anything else?
- Why did that happen? (point to Ben) Ben was sick in the morning but now he's healthy, why is that? Why is Ben healthy now?
- Why did that happen? (point to Jim) Jim was healthy in the morning and now he's still healthy, why is that? Why is Jim still healthy now?

Sick to healthy, sick to sick:

This is Jane in the morning. Jane is sick. She doesn't feel good and her body feels bad. This is Heather in the morning. Heather is sick. She doesn't feel good and her body feels bad. A whole day goes by. Now it's nighttime. This is Jane at night. Jane was sick in the morning but now she's healthy. Now she feels good and her body feels strong. This is Heather. Heather was sick in the morning and now she's still sick. She still doesn't feel good and her body feels bad.

- Why is that? Anything else?

- Why did that happen? (point to Jane) Jane was sick in the morning but now she's healthy, why is that? Why is Jane healthy now?
- Why did that happen? (point to Heather) Heather was sick in the morning and now she's still sick, why is that? Why is Heather still sick now?

Healthy to sick, healthy to healthy:

This is Melissa in the morning. Melissa is healthy. She feels good and her body feels strong. This is Felicia in the morning. Felicia is healthy. She feels good and her body feels strong. A whole day goes by. Now it's nighttime. This is Melissa at night. Melissa was healthy in the morning but now she's sick. Now she doesn't feel good and her body feels bad. This is Felicia at night. Felicia was healthy in the morning and now she's still healthy. She still feels good and her body feels strong.

- Why is that? Anything else?
- Why did that happen? (point to Melissa) Melissa was healthy in the morning but now she's sick, why is that? Why is Melissa sick now?
- Why did that happen? (point to Felicia) Felicia was healthy in the morning, now she's still healthy, why is that? Why is Felicia still healthy?

Healthy to sick, sick to sick:

This is David in the morning. David is healthy. He feels good and his body feels strong. This is Mike in the morning. Mike is sick. He doesn't feel good and his body feels bad. A whole day goes by. Now it's nighttime. This is David at night. David was healthy in the morning but now he's sick. Now he doesn't feel good and his body feels bad. This is Mike at night. Mike was sick in the morning and now he's still sick. He still doesn't feel good and his body feels bad.

- Why is that? Anything else?
- Why did that happen? (point to David) David was healthy in the morning but now he's sick, why is that? Why is David sick now?
- Why did that happen? (point to Mike) Mike was sick in the morning and now he's still sick, why is that? Why is Mike still sick now?

References

Ahn, W., Marsh, J., & Luhmann, C. (2007). Dynamic interpretations of covariation data. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 280-293). New York, NY: Oxford University Press.

Ahn, W., & Kalish, C. (2000). The role of mechanism beliefs in causal reasoning. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 199-225). Cambridge, MA: MIT Press.

Ahn, W., Kalish, C., Medin, D., Gelman, S. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition, 54*, 299-352.

Amsterlaw, J., & Wellman, H. (2006). Theories of mind in transition: A microgenetic study of the development of false belief understanding. *Journal of Cognition and Development, 7,* 139-172.

Au, T., Sidle, A., & Rollins, K. (1993). Developing an intuitive understanding of conservation and contamination: Invisible particles as a plausible mechanism. *Developmental Psychology, 29*, 286-299.

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 47-83). Oxford: Blackwell.

Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*, 946-964.

Brewer, W., Chinn, C., & Samarapungavan, A. (2000). Explanation in scientists and children. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 279-298). Cambridge, MA: MIT Press.

Bonawitz, E., Lim, S., & Schulz, L. (2007). Weighing the evidence: Children's theories of balance affect play. *Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society.* Nashville, Tennessee.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209-254). New York: Academic Press.

Callanan, M., & Oakes, L. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development, 7*, 213-233.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carey, S. (1995). On the origin of causal understanding. In D. Sperber, D. Premack, & James (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 268-308). New York: Oxford University Press.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104,* 367-405.

Cheng, P., & Novick, L. (1992). Covariation in natural causal induction. *Psychological Review, 99*, 365-382.

Chi, M., DeLeeuw, N., Chiu, M., LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439-477.

Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182.

Chin, D., Schonfeld, D., & O'Hare, L., Mayne S., Salovey, P., Showalter, D., & Cicchetti D. (1998). Elementary school-age children's developmental understanding of the causes of cancer. *Journal of Developmental and Behavioral Pediatrics, 19*, 397-403.

Chouinard, M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development, 72,* 1-112.

Cummins, R. (2000). "How does it work?" versus "What are the laws?": Two conceptions of psychological explanation. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 117-144). Cambridge, MA: MIT Press.

Frazier, B., Gelman, S.A., & Wellman, H. (under review). Preschoolers' search for explanatory information within parent-child conversation.

Gelman, S.A., & Coley, J. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26*, 796-804.

Gelman, S.A., & Markman, E. (1986). Categories and induction in young children. *Cognition, 23*, 183-209.

Gelman, S.A., & Markman, E. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532-1541.

Gelman, S.A., & Wellman, H. (1991). Insides and essence: Early understandings of the non-obvious. *Cognition, 38,* 213-244.

German, T. (1999). Children's causal reasoning: Counterfactual thinking occurs for 'negative' outcomes only. *Developmental Science, 2,* 442-447.

Glymour, C. (2000). Bayes nets as psychological models. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 169-198). Cambridge, MA: MIT Press.

Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 299-323). Cambridge, MA: MIT Press.

Gopnik, A. (1996). The scientist as child. *Philosophy of Science, 63,* 485-514.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gopnik, A., & Schulz, L. (2007). *Causal learning; Psychology, philosophy and computation.* New York: Oxford University Press

Gopnik, A., & Sobel, D. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development, 71*, 1205-1222

Gopnik, A., Sobel, D., Schulz, L., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*, 620-629.

Gopnik, A., & Wellman, H. (1994). The theory theory. In L. Hirschfeld & S.A.Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp.257-293). New York: Cambridge University Press.

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development, 75*, 409-427.

Haith, M.M. (1998). Who put the cog in infant cognition: Is rich interpretation too costly? *Infant Behavior and Development, 21*, 167-179.

Haith, M. & Benson, J. (1997). *Infant cognition*. D. Kuhn and R. Siegler (Eds.), Handbook of child psychology (5th edition), volume 2: Cognition, perception, and language development. New York: Wiley.

Harris, P., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition, 61,* 233-259.

Hempel, C. (1965). *Aspects of scientific explanation.* New York: Free Press.

Hempel. C., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15,* 135-175.

Hickling, A., & Wellman, H. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology, 37,* 668-683.

Hilton, D. (1995). Logic and language in causal explanation. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 495-529). Oxford: Clarendon Press.

Hilton, D. (1988). Images of science and commonsense explanation. In D. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 1-8). New York: New York University Press.

Hilton, D., & Slugoski, B. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review, 93*, 75-88.

Hume, D. (1777). *An enquiry concerning human understanding.* Reprint, Oxford University Press, 1975.

Inagaki, K., & Hatano, G. (2002). *Young children's naïve thinking about the biological world.* New York, NY: Psychology Press.

Jaswal, V. K., & Markman, E. M. (2007). Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development, 8*, 93-111.

Kalish, C. (1996). Causes and symptoms in preschoolers' conceptions of illness. *Child Development, 67*, 1647-1670.

Keil, F. (2006). Explanation and understanding. *Annual Review of Psychology, 57*, 227-254.

Keil, F. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Keil, F., & Wilson, R. (2000). The shadows and shallows of explanation. In F. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 87-114). Cambridge, MA: MIT Press.

Keil, F., Rozenblit, L., & Mills, C. (2004). What lies beneath? Understanding the limits of understanding. In D. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 227-249). Cambridge, MA: MIT Press.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48,* 507-531.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning.* MIT Press.

Koslowski, B., & Thompson, S. (2002). Theorizing is important, and collateral information constrains how well it is done. In P. Carruthers, S. Stich, and M. Siegal, (Eds.), *The Cognitive Basis of Science* (pp. 171-192). Cambridge University Press.

Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43*, 186-196.

Kushnir, T., & Gopnik, A. (2005). Young children infer causal strength from probabilities and interventions. *Psychological Science, 16,* 678-683.

Kemler-Nelson, D. (1995). Principle-based inferences in young children's categorization: Revisiting the impact of function on the naming of artifacts. *Cognitive Development, 10*, 347-380.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Legare, C. H., Wellman, H. M., & Gelman, S. A. Evidence for an explanation advantage in naïve biological reasoning. *Unpublished manuscript.*

Leslie, A. M., & Keeble, S. (1987). Do six-month old infants perceive causality? *Cognition, 25,* 265-288.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology, 55*, 232-257.

Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Science, 10*, 464-470.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition, 99,* 167-204.

Mackie, J. L. (1974). *The cement of the universe: A study of causation.* London: Oxford University Press.

Medin, D., Coley, J., & Storms, G. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review, 10,* 517-532.

Mills, C., & Keil, F. (2004). Knowing the limits of one's understanding: The development of an awareness of an illusion of explanatory depth. *Journal of Experimental Child Psychology, 87,* 1-32.

Murphy, G.L. (2002) The big book of concepts. MIT Press

Murphy, G.L. and Allopenna, P.D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, Cognition, 20,* 904–919.

Murphy, G., & Medin, D. (1985). The role of theory in conceptual coherence. *Psychological Review, 92*, 289-316.

Oakes, L. M., & Cohen, L.B. (1994). Infant causal perception. In C. Rovee-Collier & L.P. Lipsitt (Eds.), *Advances in infancy research, Vol 9.* Norwood, NJ: Ablex.

Onishi, K., Baillargeon, R., Leslie, A. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica, 124,* 106-128.

Piaget, J. (1929). *The child's conception of the world.* NY: Harcourt, Brace Jovanovich.

Piko, B., & Bak, J. (2006). Children's perceptions of health and illness: Images and lay concepts in preadolescence. *Health Education Research, 21*, 643-653.

Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge.* New York: Harper and Row.

Pitt, J. (1988). *Theories of explanation.* New York: Oxford University Press.

Railton, P. (1989). Explanation and metaphysical controversy. In P. Kitcher & W.C. Salmon (Eds.), *Scientific explanation: Minnesota studies in the philosophy of science, Vol. 8, Scientific.* (pp. 220-252). Minneapolis: University of Minnesota Press.

Salmon, W. (1989). *Four decades of scientific explanation.* Minneapolis, MN: University of Minnesota Press.

Salmon, W. (1984). Scientific explanation and the causal structure of the world. Princeton: Princeton University Press.

Schulz, L., & Bonawitz, E. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology, 43*, 1045-1050.

Schulz, L., & Gopnik, A. (2004). Causal learning across domains. *Developmental Psychology, 40,* 162-176.

Shanks, D. (1995). *The psychology of associative learning.* Cambridge University Press.

Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development, 47,* 1-51.

Siegler, R. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28*, 225-273.

Siegler, R., & Chen, Z. (2002). Development of rules and strategies: Balancing the old and new. *Journal of Experimental Child Psychology, 81*, 446-457.

Simon, H. (2001). "Seek and ye shall find," How curiosity engenders discovery. *Designing for Science, 5-20.*

Sloman, S.A. (1994). When explanations compete: The role of explanatory coherence on judgments of likelihood. *Cognition, 52,* 1-21.

Sloman, S.A. (2005). *Causal models: How we think about the world and its alternatives*. New York: Oxford UP.

Strevens, M. (2006). Scientific explanation. In D.M. Borchert (Ed.), *Encyclopedia of Philosophy,* MacMillan.

Tenenbaum, J., Griffiths, T., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. Special issue: Probabilistic models of cognition. *Trends in Cognitive Science, 10,* 309-318.

Trout, J.D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science, 69*, 212-233.

Trout, J.D. (2007). The psychology of scientific explanation. *Philosophy Compass, 2/3*, 564-591.

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,* 167-198.

Wellman, H. M. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In D. Kuhn & R. Siegler (Eds.), *Cognition, perception, and language* (pp. 523-573). *Vol. 2 of the Handbook of Child Psychology (5th ed.)* Editor-in-Chief: William Damon. New York: Wiley.

Wellman, H. & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43*, 337-375.

Wellman, H., Hickling, A., & Schult, C. (1997). Young children's psychological, physical, and biological explanations. In H. M. Wellman & K. Inagaki (Eds.), *The emergence of core domains of thought: reasoning about physical, psychological and biological phenomena*. San Francisco, CA: Jossey-Bass.

Wellman, H., & Lagattuta. K. (2004). Theory of mind for learning and teaching: The nature and role of explanation. *Cognitive Development, 19*, 479-497.

Wellman, H., & Liu, D. (2007). Causal reasoning as informed by the early development of explanations. In A. Gopnik & L.E. Schulz (Eds.), *Causal learning: Psychology, philosophy, computation.* NY: Oxford University Press.

Wilson, R., & Keil, F. (1998). The shadows and shallows of explanation. In *Minds and Machines, Vol 8(1) Special issue: Cognition and explanation* (pp. 137-159).

Woodward, J. (2003). *Making things happen: A theory of causal explanation.* Oxford, England: Oxford University Press.