Title

Causal attribution: insights from developmental, cross-cultural work on social and physical reasoning

Permalink

https://escholarship.org/uc/item/0vx671jb

Author

Seiver, Elizabeth

Publication Date

2013

Peer reviewed|Thesis/dissertation

Causal attribution: insights from developmental, cross-cultural work on social and physical reasoning

by

Elizabeth Anne Seiver

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

In

Psychology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Alison Gopnik, Chair
John Campbell
Tania Lombrozo
Fei Xu

Spring 2013

Causal attribution: insights from developmental, cross-cultural work on social and physical
reasoning

Copyright 2013

by

Elizabeth Anne Seiver

Abstract

Causal attribution: insights from developmental, cross-cultural work on social and physical reasoning

Elizabeth Anne Seiver
Doctor of Philosophy in Psychology

University of California, Berkeley

Alison Gopnik, Chair

This series of studies examines the relationship between causal inference and attribution from a developmental and cross-cultural perspective. In the first study, we consider how children at the ages of four and six reason about person by situation covariation information (Kelley, 1967), both younger than previously demonstrated trait biases or spontaneously using trait words. We then compare children's explanations of people's behavior to the actual behavioral evidence. Next, we extend the paradigm to include physical causation and the potential overhypotheses that guide the formation of domain specific reasoning. We will determine whether there are domain differences in perceiving virtually identical data construed as physical events or intentional actions. And finally, in the last study, we examine the role of culture in shaping the developmental trajectory of attributional style. As in physical causal inference, social causal inference combines covariational evidence and prior knowledge. We can use the tools of causal inference to understand how culturally-mediated prior beliefs affect the construal of person-by-situation information in a domain specific manner and trace the origins of attributional biases in adults.

For Mike

Acknowledgments

First and foremost I would like to thank my advisor, Alison Gopnik, for her guidance and insight throughout my graduate career. I would also like to thank the other members of my committee, Tania Lombrozo, Fei Xu, and John Campbell, for many thoughtful discussions of other facets of causal reasoning and cognition. I would also like to express my gratitude to Kaipeng Peng for facilitating and funding our research affiliated with Tsinghua University. I would also like to thank Noah Goodman and Liqi Zhu for collaborating with us on our research endeavors. I would like to thank my labmates past and present for their support over the years, especially Chris Lucas, Anna Waismeyer, and Daphna Buchsbaum, who helped this research move forward. My research would not be possible without the hundreds of parents and children both in the United States and China, who generously agreed to participate in these studies. Funding from the McDonnell Foundation and the NSF Graduate Fellowship Program made this work possible. I would also like to thank the many undergraduate research assistants who have worked hard making stimuli, collecting data, and generally being superheroes making sure everything ran smoothly. Adrienne Wente and Sophie Bridgers: working with you has been an honor and you truly inspire me.

Table of Contents

Chapter 1
Introduction

How do we explain other people's behavior? There are two main components of making attributions for why people do things. One of them is the actual data we observe in the world of different people acting in different situations. The second is our pre-existing theories about why people (or specific individuals) do what they do.

Kelley (1967) conceived of a structure that placed human behavior along several axes. One of them is the specific person performing this action. The second is the general situation that they are in. By tracking individuals across situations and how they co-occur, we can instead conceive of behaviors in terms of person and situation covariation information. This statistical approach to behavior tracking, Kelley argued, forms the basis of human attribution.

Kelley was among the first social psychologists to propose that in making an attribution, we track statistical information about the person and the situation to determine the cause of the behavior. If Mary trips while she is walking down the street, it may be because she is clumsy. If she trips several more times, you might become more confident in that attribution. If however four people in quick succession trip in the same location that Mary did, you might decide that the particular section of sidewalk is dangerous or uneven. Statistical evidence about people and situations is not an absolute measure of causality, but is an important tool to determine likely causes, and to predict future behavior (such as whether more pedestrians will trip on that stretch of sidewalk, or whether Mary will stay on her feet on the next block).

At the same time, other social psychology researchers have noted how humans are not particularly adept at tracking behavioral data. Adults have well-developed theories about the causes of people's actions, and these can sometimes have undue weight on interpreting new behavioral information, thus constituting a bias. They tend to underweight evidence that contradicts their existing theories and overweight evidence that supports their worldview. Some of these beliefs include giving oneself more credit for having circumstances lead to an undesirable action but not extending that same excuse towards other people (Jones & Nisbett, 1971; Malle, Knobe, & Nelson, 2007), and assigning more positive characteristics to in-group than out-group members (Brewer, 1979; Pettigrew, 1979). The correspondence bias, or fundamental attribution error, leads to assigning trait labels to others (especially for antisocial behavior) with scant evidence, sometimes even when it is explained away by other factors. In one of the original studies looking at this phenomenon, researchers found that adults assumed that a person who wrote a pro-Castro essay must actually be in favor of Castro (Jones & Harris, 1967). This belief persisted even when they were told that the author's position was determined by a flip of a coin.

So how do we develop these biases or theories in the first place? First, studying children at different ages gives us the opportunity to track the development of the bias to over-attribute traits to others. Social psychologists were the first to investigate children's use of trait words and compare them to adults. Even though young children do not spontaneously

describe people's behavior using trait words (Rholes & Ruble, 1984), studies have shown they can endorse them in the right situation and make behavioral predictions based on trait words (Heyman, 2009; Heyman & Gelman, 2000a; D. Liu, Gelman, & Wellman, 2007). But the issue remains if using trait language is due to an increasing understand of trait words or an increasing adoption of cultural attitudes about the fixedness of traits and the strength of their effect on people's behavior.

Even at young ages, children reason in biased ways about other people. They tend to avoid associating with other people labeled as "unlucky", even for such minor circumstances as being caught in the rain without an umbrella (Olson, Banaji, Dweck, & Spelke, 2006). A number of well-known developmental studies have shown that children are also prone to in-group and out-group biases, even when the groups are randomly assigned (Sherif, Harvey, White, Hood, & Sherif, 1961) and only distinguished by shirt color (Bigler, Jones, & Lobliner, 1997).

The correspondence bias (the more contemporary name for the fundamental attribution error) is an interesting case to consider from a developmental perspective. To readily assign a trait label where one is not warranted, you must have a familiarity with trait labels (nice, mean, generous, etc.) and presumably also use trait words on a regular basis. Children themselves do not spontaneously use trait language until the age of 7. Because of this, researchers have not studied the correspondence bias in younger children.

However, there are methods of studying an underlying preference for trait explanations that do not rely on such an explicit measure. Kelley's person by situation covariation model of attribution helps us explain why watching Mary trip multiple times would support an attribution that she is clumsy. However, it might still be "something about Mary" that is not necessarily a trait. Perhaps her shoe has broken, or she is not feeling well and so is not watching where she is walking. There are many potential causes of Mary's behavior, even when identifying the primary cause as related to her instead of the sidewalk. Traits are just one possible type of attribution that pertains to individual people.

Second, the bidirectionality of constructing theories from data and also using those theories to interpret new data is a paradigm that is well covered in the emerging literature on causal Bayes nets. Attribution specifically refers to the cause of a person's action, which falls into a broader category of cause and effect relationships. Cognitive psychologists have been studying this more general notion of causality (usually physical causality) to investigate the cues people use to infer that an object or event caused a certain outcome. More recently, researchers have begun to refine a computational model approach to studying people's beliefs about causality and larger theoretical frameworks.

This causal inference approach coalesces nicely with Kelley's approach to look at how people interpret behavioral evidence. Researchers have recently begun to take this approach into the psychological domain. Such an approach lends itself to examining the origins of attributional biases, looking at less explicit measures than language used but rather how well they are tracking person by situation covariation information.

Considering attribution from a causal inference perspective opens up new possibilities for studying the correspondence bias in children. This allows insight into children's potential bias at a younger age than would otherwise be thought possible by looking at how they track statistical information. It also allows us to explore how these biases might form in the first

place. This works not only to uncover people's underlying theories, but also models how conceptual change takes place in the form of theoretical shifts. Our experiences shape our perception of our future experiences.

The fundamental attribution error, in particular, has very clear implications from a data/theory perspective. The specific claim of the bias is that, when presented with person and situation information, we discount the situation evidence and over-rely on a dispositionalist theory. How do children form these biases? Do they display biases, and is there any difference across ages? By looking at how children track person by situation covariation information, this allows us to move beyond more explicit measures like trait words or even endorsement of trait labels.

In the following chapters, I will discuss a series of studies designed to probe the relationship between causal inference and attribution from a developmental perspective. In the next chapter, I consider how children at the ages of four and six reason about person by situation covariation information (Kelley, 1967), both younger than previously demonstrated trait biases or spontaneously using trait words, and how children's explanations of people's behavior correspond to the actual behavioral evidence. In the third chapter, I expand the scope to include physical causation and the potential overhypotheses that guide the formation of domain specific reasoning. We will determine whether there are domain differences in perceiving virtually identical data construed as physical events or intentional actions. And finally, in the fourth chapter, I examine the role of culture in shaping the developmental trajectory of attributional style.

Chapter 2
Causal inference and the development of social attribution

2.1   Introduction

People explain human actions in different ways. They may attribute a person's actions to their internal, individual, enduring characteristics or to the effect of external situations. Social psychologists have found that these causal explanations and attributions have far-reaching consequences for other kinds of social cognition and behavior, such as motivation, achievement, assigning blame, mental health, and general emotional well-being in adults (e.g. (Hong, Chiu, Dweck, Lin, & Wan, 1999; Levy & Dweck, 1998; Tamir, John, Srivastava, & Gross, 2007), and in children (Levy & Dweck, 1999; Patrick, Skinner, & Connell, 1993).

Especially in Western cultures, many adults tend to attribute the actions of others to individual, enduring traits of the person rather than to external situations (Jones & Harris, 1967; Na & Kitayama, 2011; Nisbett, 2003; Ross, Amabile, & Steinmetz, 1977). Some researchers have suggested that this is because these adults have developed an intuitive theory that explains action in terms of such traits (Molden, Plaks, & Dweck, 2006; Morris & Peng, 1994; Rosati et al., 2001). This existing theory would affect the observer's interpretation of new behavioral evidence. Just as a stubborn scientist will interpret and explain all evidence in terms of her pet theory, adults who have developed a strong prior belief that actions are the result of traits might show a bias towards trait explanations.

What kinds of evidence might lead to an attribution bias? Kelley originally suggested that reasoning from covariation evidence might play an important role in trait attributions (Kelley, 1967; Plaks, Grant, & Dweck, 2005). Empirical studies confirm that covariation — the degree to which two variables change together across contexts — plays a role in adult attribution (Cheng & Novick, 1990; Hewstone & Jaspars, 1987; Morris & Larrick, 1995; Orvis, Cunningham, & Kelley, 1975; Sutton & McClure, 2001). (See though Malle (2011) for a dissenting opinion on the person/situation dichotomy and role of covariation.)

The developmental trajectory that leads to these adult attribution biases is still unknown, however, and it is equally unknown how the role of covariation evidence changes along that trajectory. Even very young children clearly can explain actions in terms of internal psychological causes; in fact, they preferentially explain action in terms of internal mental states (Flavell, Flavell, Green, & Moses, 1990). Very young children can also understand that these mental states may differ in different individuals. For example, 18-month-olds understand that someone else may have different desires than they do (see e.g., Repacholi & Gopnik, 1997) and two-year-olds can make these differences explicit in their explanations (Bartsch & Wellman, 1995). However, traits have a more complex causal structure than simple mental states. Beyond mental states themselves, traits also possess the qualities of 1) persistent differences across different individuals and 2) consistency within a particular individual over time and across different situations. When and why do children make causal attributions of this kind?

Previous research has shown that preschool children's explanations and predictions about behavior differ from those of adults in two important ways. First, preschoolers don't spontaneously use trait words to explain actions. Many researchers have demonstrated that children do not spontaneously explain actions in terms of traits until middle childhood and that these attributions increase over time (Alvarez, Ruble, & Bolger, 2001; Higgins & Bryant, 1982; Peevers & Secord, 1973; Rholes & Ruble, 1984; Ruble, Feldman, Higgins, & Karlovac, 1979; Shimizu, 2000). Second, preschool children, unlike adults, do not spontaneously predict that an individual actor will continue to display a particular type of behavior over time or across situations. For example, when they see someone behave in a nice or mean way once, they do not predict that that pattern will continue over time or in a new context (e.g., Rholes & Ruble, 1984).

These discrepancies between children and adults might lead to the conclusion that young children simply cannot make trait attributions at all. However, more recent research has shown that children think in more 'trait-like" ways when they are given particular kinds of information. When preschoolers are shown an actor frequently exhibiting a particular behavior, they infer that the actor will continue to produce that behavior in the future (Boseovski & Lee, 2006). Similarly, if they are given a trait label (if they are told, for example, that someone is nice or mean), they can infer the sort of behaviors the person will produce (Heyman & Gelman, 2000b; 2000a; Liu et al., 2007). Conversely, if they observe many instances of a trait related behavior, they can infer the right trait label (Ferguson, Olthof, Luiten, & Rule, 1984; Heyman & Gelman, 1999; Matsunaga, 2002). In the earlier literature on consensus, preschool children were more likely to attribute a choice to the particular desire or preference of the actor when they saw many people making different choices. When different people made the same choice, they were more likely to attribute the choice to a feature of the object (Higgins & Bryant, 1982; Ruble et al., 1979).

So evidence about frequency, variation or consensus, or the use of a trait label, can influence preschoolers' attributions. However, in all these cases, the attributions might be more like simple internal mental state attributions, rather that having the distinctive features that characterize adult trait attributions. Thus hearing a trait label, or witnessing that an action was frequently produced or varied across individuals, leads young children to infer that the mental state underlying this action is frequent or variable. However, these preschoolers still did not spontaneously construct "trait-like" explanations, or use trait labels to make predictions about what different individuals would do across time or in new situations. They did not demonstrate that they interpreted trait labels as adults do, in terms of enduring and consistent features of individual people. Instead, they may have simply matched the frequency of behaviors in a particular individual, or the variation of behaviors among individuals, to relevant trait labels.

This might be because the data that children were given in these studies did not actually license the children to infer the full causal structure of adult attributions. To accurately infer internal or external causes for behavior, and predict future behavior, it is important to track multiple people across multiple situations, not just to track the frequency of behavior in a single person, or the variance of behavior across people. This richer pattern could then be more confidently generalized to a novel person or novel situation. For example, if a scientific personality psychologist wanted to claim that an action was the result of a trait,

she would have to show both that the action varied across individuals and was constant across situations – just one of these covariation patterns would be insufficient. This richer pattern of covariation, including both variation within and across individuals and variation within and across situations, would normatively support attributions with the causal structure of adult traits.

It might seem that tracking more complex covariation of this sort, and using it to infer causes, would be too difficult for very young children. Recently, however, a number of studies have shown that even very young children are surprisingly good at using covariation information to determine underlying physical causal structure, and that they do so in a rational way (Gopnik & Schulz, 2007; Gopnik et al., 2004; Gopnik, Sobel, Schulz, & Glymour, 2001). However, there is no systematic research on young children's use of this type of covariation, rather than simple frequency or variation, in the attribution and the explanation of action. Could covariation play a similar role in children's social inferences?

Such studies on covariation would extend our understanding of social cognition in young children, but they might also help us understand children's causal inference more generally. Recent developmental studies of causal inference from complex covariation have focused on somewhat narrow, specific causal attributions in the laboratory, e.g., whether a block will cause a machine to activate. They also focus on deterministic (noiseless) causal relations — in fact, there is some evidence that preschoolers assume that physical causal relationships are deterministic (Schulz & Sommerville, 2006).

Inferences about the causes of people's behavior play an important and general role in everyday life. They are also more likely to be probabilistic than purely physical causes — that is, we can never predict a person's behavior with complete certainty. Even if an observer makes trait attributions, she may not be surprised to see a timid person act bravely on occasion, but may expect that this behavior is more likely to occur in brave people. Person and situation explanations also require a fuller, more abstract causal schema rather than merely specific causal inferences. If children were to use covariation to infer traits, they would have to be capable of these more complex and general types of causal inference.

To explore these ideas, we presented four and six-year-old children with different patterns of action covariation, including probabilistic covariation, that would rationally support trait or situation inferences. We evaluated their causal explanations and predictions to determine if their attributions had a causal structure similar to the causal structure of adult trait attributions. In particular, we examined whether children attributed causes that both varied across individuals and were consistent and general within individuals. Both explanation and prediction are commonly used and valuable tools for insight into children's causal reasoning as well as social attributions, particularly when both measures are used together (see e.g., Wellman & Liu, 2007).

The tasks were designed to be developmentally appropriate for these young children. We provided a simple scenario that did not rely on interplay between multiple people, containing data points easily tracked across trials. The relevant trait, risk-taking, was less complex than characteristics previously studied in the attribution literature (such as intelligence or generosity). Risk-taking is also less heavily valenced than traits such as "nice" or "mean", and thus less confounded by general value judgments (Alvarez et al., 2001; Miller,

1984; Morris & Peng, 1994). In this respect, it also more closely parallels the case of physical causation, where valence is not an issue.

Another question concerns the developmental course of such inference. Would older children, who do spontaneously describe people in terms of traits, reason differently than younger children? Earlier studies showed that middle-school children were more likely to make trait attributions than younger children (e.g., Rholes & Ruble, 1984), in general, but there might be many reasons for this developmental pattern. Six-year-olds might simply make more accurate causal inferences from the evidence than four-year-olds given their additional experience and better information-processing abilities, and so would be more likely to accurately infer traits from behavioral data. Or four-year-olds might be biased against trait attributions and always prefer situation attributions, while six-year-olds might simply have the opposite bias and prefer trait attributions, regardless of the evidence. A third possibility is that there would be a consistent interaction between the evidence and children's prior assumptions, of the sort described in Bayesian accounts of reasoning (Griffiths & Tenenbaum, 2009). Some recent studies suggest that this kind of interaction between evidence and prior knowledge can be found in children's physical causal reasoning (Kushnir & Gopnik, 2007; Schulz, Bonawitz, & Griffiths, 2007; Sobel, Tenenbaum, & Gopnik, 2004). For example, Kushnir and Gopnik (2007) pitted children's prior belief that contact is necessary for causal interaction against the evidence that children saw. Children initially believed that a block would have to be placed on a machine to make it activate, but they were able to gradually override that belief as they gained more evidence that the block activated remotely. However, children continued to be biased towards the contact hypothesis — they were still more likely to say that the block would make the machine go when it made contact than not. In this case we might expect that six-year-olds' responses would show an interaction between the evidence and an emerging trait bias, and that four-year-olds would base their responses on the evidence. As children, at least North American children, get more evidence confirming a general "trait theory", they might develop a stronger "prior" for trait hypotheses, and require more evidence to overcome that prior. In that case, six-year-olds might actually prove to be, rather surprisingly, less sensitive to behavioral evidence than four-year-olds.

This study differs from earlier studies of covariation and trait attribution in several ways. First, we give children the equivalent of a 2 x 2 covariation table: evidence that would support explanations with the causal structure of traits, rather than simply giving evidence about marginal frequency or variation. Second, we see if children who receive these data will go beyond matching trait labels to patterns of frequency or variation and will generate spontaneous "trait-like" or "situation-like" explanations and make appropriately general predictions. Third, we include a probabilistic covariation condition to see whether and how children reason about noisy behavioral data. The current study thus integrates recent research on causal inference and on the development of social cognition.

2.2   Method

2.2.1   Participants
In the test conditions there were 48 four-year-olds (M = 4.5 years, range = 4.0-5.2 years), 26 boys and 22 girls, and 48 six-year-olds (M = 6.4 years, range = 6.0-6.9 years), 24 boys and 24 girls. The control conditions included 31 four-year-olds (M = 4.4 years, range =

4.0-4.9 years) and 32 six-year-olds (M = 6.6 years, range = 6.0-6.9 years). Recruitment and testing took place at a children's science museum and a local preschool. Although official demographic data were not collected, the participants were

2.2.2 Materials

Two small female dolls were used, as well as scaled three-dimensional colorful cardboard constructs of a diving board with swimming pool, a trampoline, and a bicycle.

2.2.3 Test Design

Person vs. Situation Conditions

In each experimental condition, participants viewed a total of eight engaging actions and eight backing away actions. However, the distribution of those actions either covaried with the situation or with the individual in a between-subjects design (see Table 1 in the appendix). We refer to the condition where behavior covaried with the dolls (but not with the situation) as the person condition, and the condition where behavior covaried with the activities (but not with the dolls) as the situation condition. In the person condition, for example, Josie would consistently play on the bicycle and the trampoline, while Sally would consistently back away from both activities. In the parallel situation condition, both Sally and Josie would consistently play on the bicycle, but they would both consistently back away from the trampoline. Importantly, everything, including language, was held constant between these conditions except for the actual covariation pattern.

Deterministic vs. Probabilistic Conditions

In the deterministic person and situation conditions, the dolls either engaged in or backed away from the activity consistently on all four trials. In the deterministic case, the experimenter referred to the appropriate mental state in narrating the events. On each trial, in both the situation and person conditions, she said either "Look, Sally/Josie's playing on the diving board/trampoline, she's not scared" or "Look Sally/Josie's not playing on the diving board/trampoline, she's scared". In the probabilistic conditions, the dolls either engaged in the activity three out of four times or backed away three out of four times. The anomalous evidence occurred on the third approach. The procedure was otherwise identical to the deterministic case, in both the person and situation condition, except that the experimenter's narration of the action was changed slightly to be more appropriate to the probabilistic context. When a scared person occasionally acts bravely, however, we tend to think that she is still scared but has overcome her fear on this occasion. So instead of saying "she's scared" the experimenter simply said, "Look! Sally's playing on the diving board" or "Look! Sally doesn't want to play on the diving board". This also meant that the mental state reference in the description of the event was much more indirectly related to the trait in this condition than in the deterministic condition.

Control Condition

In additional to the experimental conditions, we ran a control measure to obtain a comparison baseline preference for explaining behavior and to assess the potential influence of prior knowledge. In the control conditions, we tested an additional group of children to see if they would prefer a person or situation explanation when frequencies differed, but in the

absence of full 2 x 2 covariation information. One doll acted fearfully more frequently than the other in the person test condition, while each doll showed fear the same number of times in the situation test condition (see Table 1). This information alone might have caused children to make different attributions. Therefore control condition vignettes matched this frequency difference — and were otherwise completely identical to the test conditions — but did not have the 2 x 2 covariation pattern (see Table 1). In these control conditions, the evidence by itself does not rationally support a particular inference about the cause of the characters' behavior, and children's explanations should be at chance if they are based solely on this evidence. However, prior knowledge might bias children towards preferring either person or situation explanations in these cases.

In the deterministic control, the procedure was identical to the deterministic test condition except that children observed one doll play at only one activity, and the other doll play on a second activity; e.g., children watched Sally approach and play on the trampoline four times, and then saw Josie approach and back away from the bicycle four times. Although the actions and language were identical to the test conditions, in this case the covariation evidence supports both causal hypotheses equally.

However, this meant there were fewer trials for each doll overall in the control than in the test conditions, and it was unclear whether this might make the task easier or harder for the children. Therefore, for the probabilistic control, one doll approached an activity two out of eight times, and the other approached a second activity six out of eight times, mirroring both the ordering and number of positive and negative trials in the probabilistic test condition for both persons and situations (see Table 1). Again, since language and frequency were the same across the test and control conditions, this ensured that responses were not simply an effect of the linguistic descriptions or the frequency of the actions.

Test Procedure

All children were tested in a quiet room by a female experimenter and randomly assigned to each experimental condition. Participants observed a vignette in which two dolls (named Sally and Josie) made a series of approaches to two activities (chosen from among the trampoline, bicycle and diving board). The first doll would approach activity A four times, followed by the second doll, who would also approach that activity four times. Then the first doll would approach activity B four times, followed by the second doll. On each trial the doll would either engage in the activity (dive in the pool, jump on the trampoline or ride the bike), or else would back away. Doll order and activity order were counterbalanced across participants.

For example, in the deterministic person condition, children might see Josie jump on the trampoline four times, then see Sally approach the trampoline and back away four times, then see Josie ride on the bicycle four times and then see Sally approach the bicycle and back away four times. The experimenter narrated throughout on each trial as each doll either engaged in the activity or backed away (e.g. "Look! Josie's playing on the trampoline. She's not scared." in the deterministic condition or "Look! Josie's playing on the trampoline" in the probabilistic condition).

Explanation questions

At the end of the vignette, the experimenter asked two open-ended explanation questions, one for each doll's last action on the second activity (e.g., "Why did Josie jump on the trampoline?" and "Why didn't Sally jump on the trampoline?"). Children first answered for the doll most recently viewed, and then answered for the second doll. Children were free to explain the behavior as they wished. If the participant refused to answer, or gave an irrelevant answer, the experimenter would follow up with a forced-choice question contrasting a person and situation attribution (e.g., "Why did Josie jump on the trampoline? Is it because she's the kind of person who does brave things, or because the trampoline is safe to play on?" or "Why didn't Sally jump on the trampoline? Is it because she's the kind of person who gets scared, or because the trampoline is dangerous to play on?"). Children sometimes responded to the explanation questions by simply saying, "Because she wanted to/didn't want to", especially in the probabilistic condition. The experimenter followed these responses with the question "Why did she want to?" The follow-up answer was used for coding the explanation type. A scoreable response was needed before moving on to the next question.

Prediction questions

In the person and situation conditions, the experimenter then asked two prediction questions about a new future event. In each of these conditions, the evidence facilitates a particular type of novel prediction. In the person condition, the evidence allows you to make a prediction about what each doll would do in a new situation, though it does not allow you to make predictions about what a new doll would do in the earlier situations. In the situation condition, this is reversed. We tested to see if children in each condition would make the appropriate generalization and so make correct predictions. In the person condition, the participants were asked to predict what each doll would do in a new situation ("Now let's pretend that Sally and Josie go over to the diving board. Do you think Josie will play on it? <child answers> Do you think Sally will play on it?"). In the situation condition, children were asked to predict what a new doll would do in each of the earlier situations ("Now let's pretend that Sally and Josie have a friend named Mary. Do you think she'll play on the trampoline? <child answers> Do you think she'll play on the bicycle?"). Children answered either "yes" or "no." For predictions to be scored as correct, children had to answer both questions correctly. This always entailed one "yes" and one "no" answer since playing and not-playing were contrasted in both conditions. That is, children in the person condition had to say that each doll would act consistently (and differently from each other) in the new situation and children in the situation condition had to respond that a new doll would act consistently in (and differentiate between) each of the old situations. All other patterns of responses were scored as incorrect. Thus if children were responding at chance, they would be scored correct 25% of the time.

Explanation coding

As noted above, if children did not provide a relevant explanation spontaneously, they received the forced-choice question. Both forced-choice and open-ended explanation responses were coded by observers into two mutually exclusive response types (k, inter-rater reliability = .732, $p <.001$). The observers were blind to study condition and only saw the explanations themselves. Therefore differences across conditions would suggest that the

coding scheme had some validity as well as reliability. (For examples of both types of explanations, see Table 2.) An explanation was coded as a "person" response if it attributed the doll's behavior to an internal cause specific to that doll and in contrast to others, similar to the Ruble et al (1979) construct of "person attribution". This cause could involve the doll's mental states, such as consistent desires or beliefs, or refer to other stable characteristics of the person, such as personality, age or size. This category included classic "trait" attributions, such as "she's brave" but also included a wider variety of person-specific attributions such as "she's the big sister" or "she knows how to ride a bike" or "she likes to swim". Thus trait attributions are subsumed into the broader category of "person attributions."

An explanation was coded as a "situation" response if it referred to an underlying cause that was outside the doll. This included both aspects of the physical situation (bounciness of the trampoline, pool temperature, etc.) and the social situation ("she did it because her friend did it"). All the spontaneous explanations could be coded into one of the two categories. Finally, forced choice responses were coded "person" or "situation" based on which of the two options was chosen.

Some person and situation explanations might seem similar at first glance; for example, several children drew upon their own recent experiences about learning to ride a two-wheeler bike for both person explanations ("she doesn't know how to ride a bike") and situation explanations ("the bike only has two wheels"). However, interestingly, as we will see below, the prevalence of each explanation type consistently varied by study condition. Children chose to stress the person or situation differentially in their explanations, often as the subject of the sentence, even when the overall content of the explanations was similar. That would suggest that the coding captured genuine differences in attribution. When they faced different patterns of covariation, children produced different types of explanations – explanations that blind observers reliably classified as stressing persons or situations.

As noted above, children were asked two explanation questions, one about each doll. They were given a "person" score of zero, one or two depending on how many "person explanations" they provided.

2.3   Results

2.3.1   Explanation results

Across study conditions, four-year-olds provided 48 forced choice responses and 106 explanations, and six-year-olds provided 12 forced choice responses and 148 explanations.

Consistent with earlier studies, only a few children gave "classic" personality trait explanations, such as "she's brave". Importantly, however, all the children who gave two person explanations not only referred to internal states of the person, but also differentiated between the two actors (e.g. they said Josie likes to swim, Sally doesn't like to swim). Moreover, both mental state and non–mental state person attributions often, though not always, implied some enduring feature, e.g., "she likes swimming", or "she's old". Examples of these explanations are in the appendix.

We began by conducting a 2 X 2 X 3 (age (4 vs. 6) X consistency (deterministic vs. probabilistic) X condition (person vs. situation vs. control) ANOVA on the person scores. A parallel ANOVA that excluded the forced-choice responses yielded the same pattern of results

as the analysis that included them, suggesting that the results were not due to the particular question format. Therefore we combined the two question types in our analyses.

As shown in Figure 2.1, there was a main effect of condition ($F(2, 153) = 19.242$, $p < .001$, $\eta^2 = .19$ overall, $F(2, 76) = 15.05$, $p < .001$, $\eta^2 = .28$ for four-year-olds; $F(2, 77) = 5.82$, $p < .01$, $\eta^2 = .13$ for six-year-olds), with more person responses in the person condition ($M = 1.73$, $SD = .494$) than in the control condition ($M = 1.32$, $SD = .758$) and than in the situation condition ($M = .85$, $SD = .825$). There was also a main effect of age; six-year-olds gave more person-based explanations than four-year-olds across conditions ($F(1, 147) = 9.87$, $p < .01$, $\eta^2 = .05$). There was no effect of the deterministic vs. probabilistic condition, (probabilistic: $M = 1.55$, $SD = .67$; deterministic: $M = 1.37$, $SD = .76$) and there were no interaction effects.

Since there were neither significant main effects nor significant interactions involving the deterministic and probabilistic conditions, these two groups were collapsed for subsequent analyses. Note that there were also minor differences in the deterministic and probabilistic condition procedures as noted above – in particular, the probabilistic test condition contained fewer references to mental states, and the probabilistic control condition used eight rather than four examples of each behavior. These differences appeared to have no effect on the children's responses. (Non-parametric analyses yielded an identical pattern of results to those given by the ANOVAs.)

Closer examination of the data showed that the four-year-olds were actually more accurate than the six-year-olds in the situation and control conditions. In the person condition, both age groups scored near ceiling, significantly different from chance ($t(23) = 6.3$, $p < .001$, $d = 2.63$ for four-year-olds; $t(23) = 8.3$, $p < .001$, $d = 3.46$ for six-year-olds) and not significantly different from each other ($F(1, 46) = .084$, $p > .5$). In contrast, in the situation condition, six-year-olds gave significantly more person explanations than four-year-olds overall ($F(1, 70) = 5.7$, $p < .05$, $d = .11$). Four-year-olds gave significantly fewer person explanations in the situation condition than expected by chance ($t(23) = -2.8$, $p < .01$, $d = -1.17$), six-year-olds were at chance ($t(23) = .720$, $p > .4$). In addition, the distribution of their scores was at chance as revealed in a goodness-of-fit test (c^2 (2, N = 24) = 4.92, $p > .08$) indicating that there were no consistent patterns within children; children were no more likely than chance to give either two person or two situation explanations.

The control conditions were identical in language and behavioral frequency to the respective test conditions, but the covariation information did not favor a person or situation attribution. Four-year-olds were at chance in the controls ($t(30) = .441$, $p > .5$). Six-year-olds, however, displayed a significant preference for person explanations ($t(31) = 5.14$, $p < .001$, $d = 1.85$). Both four-year-olds ($t(53) = 2.29$, $p < .05$, $d = .63$) and six-year-olds ($t(54) = 2.23$, $p < .05$, $d = .61$) produced significantly fewer person explanations in the situation condition than in the controls. For four-year-olds, the person and control explanations also differed significantly ($t(53) = 3.33$, $p < .01$, $d = .91$), but the difference for six-year-olds did not reach significance ($t(54) = 1.26$, $p > .2$).

Overall, as Figure 2.1 shows, both four- and six-year-olds are sensitive to the pattern of covariation when deciding what kind of causal explanation to provide – they provide more "person" explanations in the person condition than in the control condition, and more in the control than in the situation condition. However, six-year-olds, unlike four-year-olds, have a

consistent bias towards person explanations. This means that in the situation and control conditions, four-year-olds were actually more accurate, in strictly covariational terms, than six-year-olds.

Figure 2.1: Four- and six-year-olds' person responses in each covariation condition.

Error bars: 95% CI
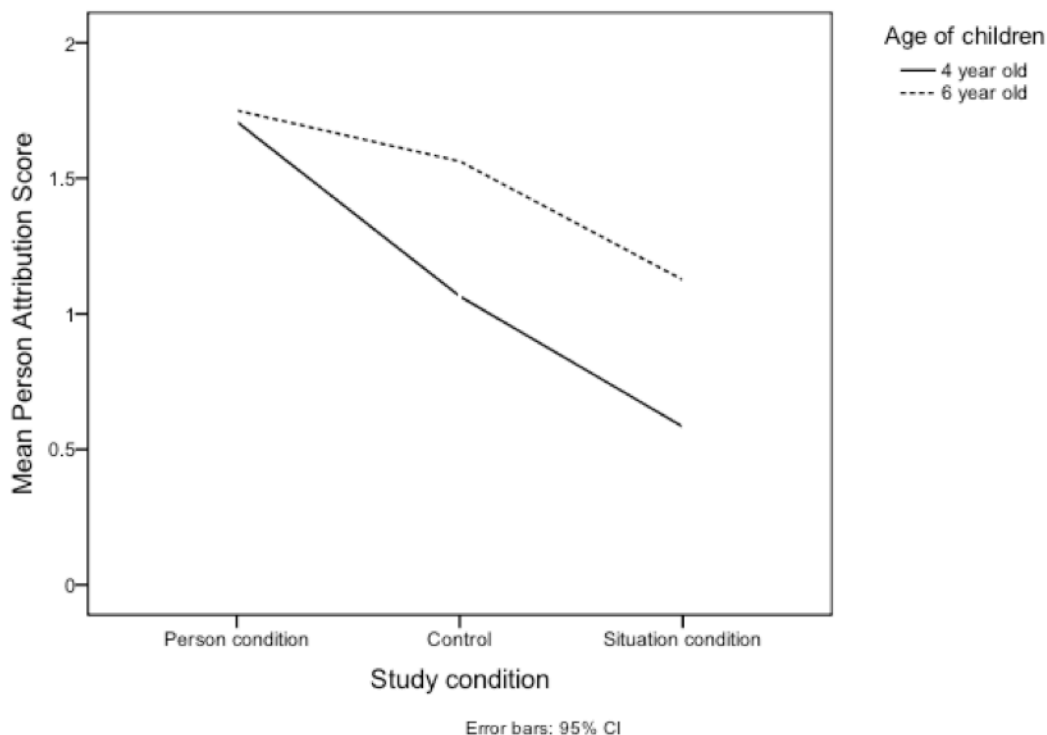
2.3.2   Prediction Results

The prediction task was designed to test whether children had only inferred a single causal explanation from the data they had actually seen, or if they had inferred a more abstract causal schema. Trait attributions not only involve causes that are internal to the actor and contrast with other actors, but also imply that these internal causes will lead the actor to behave in similar ways across situations and through time. Similarly, situation attributions not only imply that this actor behaved as she did because of the situation, but that other actors in similar situations will behave in the same way.

If children in the person condition had inferred this more abstract "trait-like" causal scheme, then they ought to predict that each doll would behave consistently in a new situation. For example, if Josie failed to approach both the bicycle and the trampoline she should also be reluctant to dive off the board. In fact, in the person condition, more four- and six-year-olds predicted that both dolls' behavior would be consistent in a novel situation than would be predicted by chance (17 out of 24 6-year-olds, *p* <.001, Binomial test; 20 out of 24

4-year-olds, $p <.001$, Binomial test; 37 out of 48 participants, $p <.0001$, Binomial test). (Note that chance here was 25% since to be scored as correct, children had to respond to two forced-choice questions.)

Similarly, in the situation condition, children who genuinely made a situation attribution should infer that a new person (Mary) would behave similarly to Josie and Sally in the two situations. If Josie and Sally both jumped on the trampoline but avoided the diving board, then so should Mary. In fact, four-year-olds tended to predict that the new doll would behave as the other dolls had done (12 out of 24 participants, $p <.01$, Binomial test) but six-year-olds did not (9 out of 24 participants, $p > .05$, Binomial test) — they tended to assume Mary would behave the same way across both situations. These results are consistent with the explanation results. They suggest that four-year-olds may generalize from covariation information rationally to make new predictions about both persons and situations, but that six-year-olds are more likely to make generalizations about persons than situations.

2.4   Discussion

In this study, four and six-year-old children used covariation information to make corresponding inferences about the causes of human actions. Given the appropriate covariation-by-person evidence, even with probabilistic data, four-year-olds explained actions in terms of internal, individual and enduring causes. They also made appropriate predictions about an individual's behavior in a new situation. When covariation evidence supported a situation attribution, they would also make those attributions and predictions correctly. Six-year-olds also used covariation information to explain and predict, but in contrast to the younger children (and more like adults), they showed an overall bias for person explanations over situation explanations. This bias apparently led them to place less weight on the behavioral evidence that was presented in the vignettes.

In particular, the presence of appropriate covariation information leads even four-year-old children to spontaneously infer causes with the full causal structure of traits, to use these causes to explain behavior, and to predict future behavior. This was true even though the only difference across conditions was the covariation pattern — language and other cues were held constant. Note also that children made these inferences significantly less frequently in the control conditions where frequency information was available, but the full covariation matrix was not.

In particular, four-year-olds invoked a "trait-like" causal schema to generate consistent predictions about people in novel situations. There are two ways of thinking about this schema. Four-year-olds may already have a trait-like schema in place, but unlike adults, initially they may apply it only in restricted conditions; when it is explicitly described by a trait label, as in the Heyman & Gelman (1999) study, or when it is strongly supported by covariation information, as in the present study.

Alternatively, the covariation information, along with other evidence from everyday life, may actually lead the children to posit a "trait-like" schema. Even though the children did not attribute full-blown traits in the same way that adults do, they did explain actions in terms of enduring, consistent but individually variable internal causes, and their predictions revealed similar causal attributions. The fact that children seemed to spontaneously invent trait-like explanations (e.g., "she's bigger", "she knows how to ride a bicycle") may support this idea.

Children may spontaneously invent such causal schemas to explain covariation patterns in particular cases, and then generalize those schemas. This kind of inference fits the pattern of general schema inference described by Kemp, Goodman and Tenenbaum (2008).

In contrast, for the six-year-olds, this schema may have been confirmed by covariation data many times and across many contexts and become an entrenched intuitive theoretical framework. Explanatory hypotheses that fit this framework would receive a higher probability at the outset, and require more data to defeat them. This sort of default theoretical framework could lead to a person bias.

What kinds of evidence could lead to this developmental change? One interesting hypothesis is that the developments at six are related to the increase in peer group interaction in middle childhood. In peer interaction, individual traits, rather than social roles or situations, will account for much of the variance in behavior. In a classroom of 20 otherwise similar children placed in a similar situation on the playground, some will consistently take risks and others will not. Children will see more trait-based covariation as they pay increasing attention to their peers, and acquire rich data sets across individuals and situations to draw upon.

Similarly, cross-cultural differences in covariation evidence may influence the development of attribution. Miller (1984) suggested that children across cultures began with similar attribution patterns and then diverged towards the more extreme adult patterns as they grew older, a claim which has been supported by further studies with children (Gonzalez, Zosuls, & Ruble, 2010; Kalish, 2002; Lockhart, Nakashima, Inagaki, & Keil, 2008). Again, these results suggest a mechanism by which cultural differences may influence the course of attribution. This may either be because members of different cultures actually do behave differently, or more probably, because culture and experience influence the information children receive from adults about traits, such as adult trait language. This evidence is especially relevant to the development of causal schemata. If people within a culture tend to describe behavior in terms of traits, then this will lead to covariation between certain behaviors and trait labels, which might itself provide evidence for a trait-schema (see Kemp et al., 2008). If children are using covariation information about people's behavior and adult trait language to infer both specific causes and more general causal schemas, such differences in the data could affect their adult social cognition. It would be very interesting to see if children in a less trait-based culture (such as mainland China) would show a similar pattern of results. One might predict that in such a culture four-year-olds would show a similar pattern, but six-year-olds would not manifest the same trait bias. We are currently conducting such studies.

These results are also interesting because they may point to broader mechanisms for learning about traits and situations. Recent computational work outlines how attributional learning might take place. In particular, causal Bayes net learning mechanisms (Pearl, 2000; Spirtes, Glymour, & Scheines, 2000) can be used to model causal reasoning and learning in adults (e.g., chapters in Gopnik & Schulz, 2007; Rehder & Hastie, 2001; Steyvers, 2003; Waldmann & Martignon, 1998), preschool children (e.g. Gopnik et al., 2001; 2004), and even infants (Sobel & Kirkham, 2006). These models predict which causal inferences should rationally be made from different patterns of covariation and prior knowledge.

Bayesian models of causal learning theories (e.g., Griffiths & Tenenbaum, 2009), in particular, suggest that children make new inferences by systematically combining prior knowledge and current covariation evidence to arrive at the right causal hypothesis. Learners

can select hypotheses rationally in the light of data by using Bayes' rule to combine the prior probability of different causal hypotheses and the probability of the current evidence given each hypothesis. Several recent studies (Kushnir & Gopnik, 2007; Schulz et al., 2007; Sobel et al., 2004) suggest that preschoolers can combine prior knowledge with covariation evidence in this Bayesian way. Moreover, recent work shows that this kind of inference can be used not only to develop specific causal hypotheses but also to construct more abstract causal schemas or "framework theories" (Griffiths & Tenenbaum, 2005; Kemp et al., 2008; Schulz et al., 2007).

This suggests a potential mechanism for the development of attribution. Children may begin by forming theories based on both people's behavior and how adults explain such behavior. They continue to overweight the evidence that confirms a culturally-conferred hypothesis or abstract causal schema, particularly the hypothesis that internal traits cause actions, while underweighting the evidence that contradicts this hypothesis. Once that schema has been highly confirmed, it will be more difficult to overturn in future, though it might still be overturned with sufficient evidence. Eventually, in adulthood, this may result in a consistent "trait bias" that is difficult to overcome.

Whether or not this account of how children learn this bias is correct, the current study shows that some of the prerequisites for such an account are in place. Children as young as four can use covariation evidence to make behavioral attributions, and six-year-olds combine that evidence with prior biases to arrive at similar (but slightly skewed) conclusions. This mirrors children's ability to infer causes in the physical domain using both prior knowledge and evidence. Further research is needed to explore a potential broader underlying framework of causal inference connecting the social and physical domains. Nonetheless, we can see the origins of Kelley's (1967) social schemata even in preschoolers.

Chapter 3 Children's causal reasoning: domain specificity

3.1   Introduction

The previous chapter shows that children's attributions for other people's behavior confirm Kelley's (1967) person by situation covariation schema. However, the primary developmental shift from age four to six does not show increased attention to the data or more sophisticated statistical understanding, but a decreasing sensitivity to the data. Specifically, we found that the older children have a preference for internal explanations about the person that interacts with the overall covariation pattern with which they are presented. This suggests that as children are getting older, they are developing stronger prior beliefs about the causes of people's behavior.

Kelley's claim about the pre-eminence of tracking person by situation covariation information cannot account for this developing attributional bias. However, this finding has precedence in the adult social psychology literature on attributional biases. When adults commit the fundamental attribution error (e.g. Jones & Harris, 1967; Ross, 1977), they assume people's behavior is due to internal (and stable) causes, even with little or no evidence to support this stance. It appears that six-year-olds' performance may reflect a shift to that type of biased thinking.

Some social psychologists have considered what the ramifications are of this well-documented attributional bias for reasoning in other domains. Morris and Peng (1994) and Peng and Knowles (2003) have explored specifically how social reasoning might affect physical reasoning. Although American adults have a standard interpretation of agents and inanimate objects in traditional Michottean launching tasks (Morris & Peng, 1994) there is some evidence that they tend to construe other aspects of Newtonian mechanics from a skewed perspective that favors dispositional explanations (Peng & Knowles, 2003). For example, they may believe that an object floating on water has an inherent "buoyant" property, as opposed to considering the relative densities of the object and the water. The idea of an object in a certain context roughly maps on to the idea of a person in a certain situation.

Other social psychologists (e.g., Nisbett, Peng, Choi, & Norenzayan, 2001) have placed the fundamental attribution error (more recently referred to as the correspondence bias) in an overarching theoretical framework that spans both psychological and physical causes. For American culture, this framework is analytic (as opposed to holistic). In analytical reasoning, the focus is on categorization, rules, and formal logic, and tracking qualities of individual objects. In holistic cognition, the causal area is more inclusive and expansive, and the focus is on the set of relationships between multiple objects. In this view, attribution is part of a larger theoretical stance that favors dispositional explanations across domains within an analytical framework. These claims cross over from social psychology into cognitive psychology when describing domain-general theories that influence judgements about both objects and people.

As was discussed in the previous chapter, Kelley's covariation model fits within the larger framework of causal Bayes nets. This goes a step beyond immediate causal inferences, and can address the issue of the formation and prevalence of emerging attributional biases,

and the formation of prior beliefs that influence the intake of future data. It also addresses the idea of overhypotheses (Goodman, 1955; Kemp, Perfors, & Tenenbaum, 2007; Lucas, Gopnik, & Griffiths, 2010) and meta-theories (that is, how theories about objects or people might be formed in the first place), such as with domain-general or domain-specific learning mechanisms.

The holistic/analytic theory connecting attribution to physical domains is a domain-general meta-theory about learning. This theory and its competitors can be encapsulated by causal Bayes nets. For example, a domain-specific meta-theory would posit that we form expectations about the physical and psychological domains separately (though this would not necessarily preclude one from influencing the other). In cognitive development research, this debate over whether we reason in domain specific or domain general ways has been largely conflated with the debate between nativism and empiricism. Nativists have argued that we are born with rudimentary core knowledge of the world with different expectations for physical objects and agents (Baillargeon, 2008; Spelke, 1994; Spelke, Breinlinger, Macomber, & Jacobson, 1992). In the even more extreme modularity view (Scholl & Leslie, 2001), domain specific reasoning is thought to involve completely separate and highly specialized modules. Empiricists, on the other hand, believe that knowledge, and therefore domains, are constructed. These researchers have emphasized the role of learning (Gibson, 1995; Karmiloff-Smith, 1994; Rumelhart & McClelland, 1987; Thelen & Smith, 1994), the extreme position of which would be the concept of tabula rasa. Empiricists would argue that while we learn from our experiences in the world, this process is the same for all learning and any domain differences are constructed.

Importantly for this study, there is a difference between domain specificity and innateness (R. Gelman, 2000). Wellman and Gelman (1992) introduced the idea of framework theories to cognitive development, which correspond to the idea of meta-theories and how we might form domain-specific expectations via experience. They attempted to reconcile the nativist concept of core domains with the empiricist emphasis on learning, suggesting that people start with separate frameworks for reasoning about different domains. These frameworks cohere into more developed theories via a general causal inference mechanism. Causal inference has provided a balanced alternative to the two extremes: domain-specific concepts might be learned. S. A. Gelman and Noles (2010) further develop the idea of learned domain-specific theories and offer two possible theoretical viewpoints on domain specificity from the perspective of causal inference: children may simply accrue greater amounts of knowledge over time – a view emphasizing continuity, or they may undergo fundamental conceptual restructuring and reorganization, leading to wholly new concepts that are incommensurate with old concepts – a view positing discontinuous stages of development.

Some developmental work has already found evidence of children reasoning in domain-specific ways from novel data and taking contextual information into account when evaluating covariation information. After all, many of the mechanisms in psychological causes do not apply to regular physical causes: mental states, beliefs, thoughts, desires. In some cases, children will not cross domains in the absence of such a mechanism, even with very strong covariation evidence. And when two events without a plausible mechanism connecting them occur with spatial and temporal contiguity, young children will not think that there is a causal relationship between them (Shultz, 1982). Children will also override within-

domain covariation evidence when there is not a plausible mechanism. If they see one light when pressed always turns on and another light when pressed never works, if the batteries are transferred from the working to the non-working light, even three-year-olds will pick the light that never worked before as the one now most likely to turn on (Buchanan & Sobel, 2011).

Other studies have shown that the domain that evidence is presented in also affects the construal of covariation data, even in children. For example, Sobel and Munro (2009) conducted a comparison study where four-year-old children were shown blocks placed on a "blicket" detector, some of which activated the machine. In one condition, it was presented as a blicket machine that detected blickets (a special category of blocks), and in the other, the machine was anthropomorphized as Mr. Blicket who liked blickets and disliked the other blocks. The groups viewed identical statistical evidence of blocks turning on the machine. Children were more likely to judge that a previously unseen internal mechanism caused Mr. Blicket to like them than caused the plain machine to activate.

Let us consider the Nisbett meta-theory as it contrasts with the framework meta-theory. While adults might have domain-specific beliefs about the relationships between causes and effects, they may interpret the covariation information, or the relationship between a within-domain cause and effect, in the same manner across domains. One possibility is that reasoning about the psychological domain affects how we construe physical events. If children consider mechanism information when forming judgments of covariation data, we would not expect an internal bias when describing physical events. This raises the question of whether the "internal bias" previously observed in six-year-olds when explaining people's behavior also maps onto the physical domain and considers the dispositional properties of objects. This would imply that children would use statistical evidence about the person and situation (as well as object and context) in a biased manner.

If we are constructing framework theories about different domains, this raises the question of the timeline of their development. At which age would we have evidence that they are reasoning in a domain-specific way, and what would the evidence for that look like? This also must be reconciled with the attribution work with adults showing issues with crossing domains. At what age could we detect an example of an "internal bias" in physical reasoning similar to the correspondence bias? Children who do not yet show an internal bias for attribution would be unlikely to have an internal bias in the physical domain. However, at some point in developing an internally-biased framework theory about people's behavior, this may be either a domain-specific overhypothesis or a more general trend towards implicating dispositional properties of both people and objects.

So which is it for young children, a domain specific view favoring qualities of people over situational influences (supported by framework theories of development), or a domain general meta-theory where we learn about domains in the same manner? Peng and Knowles (2003) directly linked object and context to person and situation, but they did not choose perfectly analogous physical events to correspond to psychological events. To best evaluate this claim developmentally, we should compare explanations for physical and psychological causes in nearly identical scenarios save for domain information. We decided to take our psychological findings with children into the physical domain in the closest possible way. What if we exposed children to nearly identical visual information, of dolls "approaching" toys and then using them, but instead referred to them as objects instead of people? Now

instead of Sally and Josie playing on the trampoline, the Sally doll and Josie doll stick to the trampoline. Instead of tracking people's behaviors across situations, now children track interactions across two categories of physical objects. And while this stays rooted in the physical domain, it still retains anthropomorphic qualities from the previous version.

In the current chapter's study, we set out to answer several questions. Will six-year-olds rely more heavily prior beliefs than four-year-olds, and will the bias carry over from the psychological domain to the physical domain as suggested in Peng and Knowles (2003)? Do children whose explanations are congruent with the pattern of data also predict that future events will retain the consistent pattern? And finally, will we find evidence for the formation of domain-specific or domain-general theories between the ages of four and six?

3.2   Method

3.2.1   Participants
160 children participated in the experiment. In the test conditions there were 48 four-year-olds (M = 4.48 years, range = 3.85-4.98 years), 23 boys and 25 girls, and 48 six-year-olds (M = 6.47 years, range = 4.98-7.19 years), 29 boys and 19 girls. The control conditions included 32 four-year-olds (M = 4.49 years, range = 4.00-4.98 years), 16 boys and 16 girls, and 32 six-year-olds (M = 6.51 years, range = 5.98-7.05 years), 20 boys and 12 girls. Recruitment and testing took place at a children's science museum and local preschools. Although official demographic data were not collected, the participants were representative of the community in the surrounding area.

3.2.2   Materials
Two small female and one male dolls were used, as well as scaled toys of a scooter, skateboard, and vault springboard. Two out of three of each had magnets surreptitiously placed on them. All of the standing surfaces of the toys and the bottoms of the dolls' feet were covered with colored construction paper or tape to obscure any visual cues about which ones were actually 'sticky'.

3.2.3   Test Design

Doll vs. Toy Conditions
In each experimental condition, participants observed a total of eight "sticking" actions and eight "bouncing away" actions. However, the distribution of those actions either covaried with the toy or with the doll in a between-subjects design (see Table 1). We refer to the condition where behavior covaried with the dolls (but not with the toy) as the doll condition, and the condition where behavior covaried with the toys (but not with the dolls) as the toy condition. In the doll condition, for example, the Josie doll would consistently stick to the scooter and the skateboard, while the Sally doll would consistently bounce away from both toys. In the parallel toy condition, both Sally and Josie would consistently stick to the scooter, but they would both consistently bounce away from the skateboard. Importantly, everything, including language, was held constant between these conditions except for the actual covariation pattern.

Deterministic vs. Probabilistic Conditions

In the deterministic doll and toy conditions, the dolls either stuck to or bounced away from the toy consistently on all four trials. On each trial, in both the situation and doll conditions, she commented on the doll either sticking (e.g., "Look, the Josie doll is sticking to the trampoline, it's not bouncing away" or not sticking (e.g., "Look the Sally doll is not sticking to the bicycle, it bounced away)". In the probabilistic conditions, the dolls either stuck to the toy three out of four times or bounced away three out of four times. The anomalous evidence occurred on the third approach. The procedure was otherwise identical to the deterministic case, in both the doll and toy condition.

Control Condition

In additional to the experimental conditions, we ran a control measure to obtain a comparison baseline preference for explaining "stickiness" and to assess the potential influence of prior knowledge. In the control conditions, we tested an additional group of children to see if they would prefer a doll or toy explanation when frequencies differed, but in the absence of full 2 x 2 covariation information. One doll would bounce away more frequently than the other in the doll test condition, while each doll stuck the same number of times in the situation test condition (see Table 1). This information alone might have caused children to make different attributions. Therefore control condition vignettes matched this frequency difference — and were otherwise completely identical to the test conditions — but did not have the 2 x 2 covariation pattern (see Table 1). In these control conditions, the evidence by itself does not rationally support a particular inference about the cause of the dolls' stickiness, and children's explanations should be at chance if they are based solely on this evidence. However, prior knowledge might bias children towards preferring either doll or toy explanations in these cases.

In the deterministic control, the procedure was identical to the deterministic test condition except that children observed one doll paired with only one toy, and the other doll stick to a second toy; e.g., children saw the Sally doll approach and stick to the skateboard four times, and then saw Josie approach and bounce away from the scooter four times. Although the actions and language were identical to the test conditions, in this case the covariation evidence supports both causal hypotheses equally.

However, this meant there were fewer trials for each doll overall in the control than in the test conditions, and it was unclear whether this might make the task easier or harder for the children. Therefore, for the probabilistic control, one doll approached a toy two out of eight times, and the other approached a second toy six out of eight times, mirroring both the ordering and number of positive and negative trials in the probabilistic test condition for both dolls and toys (see Table 1 in the appendix). Again, since language and frequency were the same across the test and control conditions, this ensured that responses were not simply an effect of the linguistic descriptions or the frequency of the actions.

Test Procedure

All children were tested in a quiet room by a female experimenter and randomly assigned to each experimental condition. The experimenter first asked participants about their experience with magnets or magnet-like toys. "Do you know how some toys stick together like this" <hands together> "but can sometimes bounce away from each other?" <hands

apart> "Well I have some toys here that can do that, and I want to see what sticks to what. Can you help me do that?" Participants then observed a vignette in which two dolls (from the set of three named Sally, Josie, and Bobby) made a series of approaches to two toys (chosen from among the skateboard, scooter and vaulting springboard). To keep the procedure as consistent as possible with the psychological case, the experimenter moved the doll as if it were a person approaching an activity. The experimenter would have the first doll approach toy A four times, followed by the second doll, who would also approach that toy four times. Then the first doll would approach toy B four times, followed by the second doll. On each trial the doll would either stick to the toy, whereupon the experimenter would move the doll and toy as if it were playing. Otherwise, the doll would initially approach and "land" on the toy, but would then bounce away (replicating the case where the character backed away in fear). Especially for the probabilistic condition, the experimenter sometimes had to "fake" sticking or bouncing, regardless of actual magnets present. Doll order and toy order were counterbalanced across participants.

For example, in the deterministic doll condition, children might see Josie stick to the scooter four times, then see Sally approach the scooter and bounce away four times, then see Josie stick to the skateboard four times and then see Sally approach the skateboard and bounce away four times. The experimenter narrated throughout on each trial as each doll either stuck to the toy or bounced away (e.g. "Look! The Josie doll is sticking to the scooter. She's not bouncing away.") The language subtly differed from the psychological case in that instead of being called by their names and with an appropriately gendered pronoun, the dolls were called "the Sally doll" etc and were referred to by the pronoun "it".

Explanation questions

At the end of the vignette, the experimenter asked two open-ended explanation questions, one for each doll's last action on the second toy (e.g., "Why did Josie stick to the trampoline?" and "Why didn't Sally stick to the trampoline?"). Children first answered for the doll most recently viewed, and then answered for the second doll. Children were free to explain the behavior as they wished. If the participant refused to answer, or gave an irrelevant answer, the experimenter would follow up with a forced-choice question contrasting a doll and toy "stickiness" attribution (e.g., "Why did the Josie doll stick to the scooter? Is it because she's the kind of doll that sticks to things, or because the scooter is sticky and pulls things in?" or "Why didn't the Sally doll stick to the scooter? Is it because she's the kind of doll who doesn't stick to things, or because the trampoline [isn't sticky enough]?"). A scoreable response was needed before moving on to the next question.

Prediction questions

In the doll and toy conditions, the experimenter then asked two prediction questions about a new future event. In each of these conditions, the evidence facilitates a particular type of novel prediction. In the doll condition, the evidence allows you to make a prediction about what each doll would do on a new toy, though it does not allow you to make predictions about what a new doll would do on the previously viewed toys. In the toy condition, this is reversed. We tested whether children in each condition would make the appropriate generalization and so make correct predictions. In the doll condition, the participants were asked to predict what each doll would do with a new toy ("Now here's my other toy, the bouncy board. Do you

think the Josie doll will stick to it? <child answers> Do you think the Sally doll will stick to it?"). In the toy condition, children were asked to predict what a new doll would do with each of the earlier toys ("Now here's my other doll, the Bobby doll. Do you think he'll stick to the scooter? <child answers> Do you think he'll stick to the skateboard?"). Children answered either "yes" or "no." For predictions to be scored as correct, children had to answer both questions correctly. This always entailed one "yes" and one "no" answer since sticking and not-sticking were contrasted in both conditions. That is, children in the doll condition had to say that each doll would act consistently (and differently from each other) in the new situation and children in the toy condition had to respond that a new doll would interact consistently with (and differently between) each of the old toys. All other patterns of responses were scored as incorrect. Thus if children were responding at chance, they would be scored correct 25% of the time.

Explanation coding

As noted above, if children did not provide a relevant explanation spontaneously, they received the forced-choice question. Both forced-choice and open-ended explanation responses were coded by observers into four mutually exclusive response types (k, inter-rater reliability = .878, p <.05). The observers were blind to study condition and only saw the explanations themselves. Therefore differences across conditions would suggest that the coding scheme had some validity as well as reliability. An explanation was coded as an "internal" response if it attributed the stickiness to the doll. This cause could involve the doll's shoes, or refer to other characteristics such as stickiness as in adhesiveness, or attraction as in magnets, or metal, or what the doll was made from. (This corresponds to a 'person' attribution in the 2012 paper.) An explanation was coded as an "external" response if it referred to an underlying cause that was outside the doll (stickiness of the trampoline, etc.). (This corresponds to a 'situation' explanation in the 2012 paper.) An explanation was coded as an "interaction" response if it explicitly cited both the doll and the toy as being part of the sticking. They could implicate a wide range of features. For example, some children thought the doll had metal feet and the scooter had a magnet, whereas other children thought the scooter was metal and the doll had magnetic feet. Any explanation not codeable into those three categories was "Other". Finally, forced choice responses were coded "internal" or "external" based on which of the two options was chosen.

3.3   Results

We recoded explanation data previously published in Seiver, Gopnik & Goodman (2012) to fit with the new coding scheme used with the magnet data. Previously termed "person" (internal) and "situation" (external) explanations, we recoded them to include Interaction and Other explanation categories. We refer to this previously-run experiment as occurring in the Psychological (person) domain, because the dolls are treated like people, with mental states and motivations. The dolls are simply stand-ins for actual people, or tools of pretense. In the Physical (magnet) domain, the physical dolls themselves are now the object in question.

We performed two kinds of analyses to examine the explanation data. Each child provided two explanations that were coded into four categories. Our first quantitative

approach was to assign four scores to each child, one for each explanation type. For internal, external, interaction, and other, a participant would receive a score of 0, 1, or 2 depending on the number of explanations given for each type. Therefore each individual participant would receive four scores that would always sum to 2. (The means of each score are presented in Table 1.)

A second quantitative approach focused more closely on the relationship between explanation type and covariation information. We calculated congruence scores, with two different potential values – congruent or incongruent – indicating whether the two explanations matched with the covariation information. In this case, we excluded the Control condition.

To examine the consistency of children's explanations, we first performed a 4 x 4 Chi Square comparing children's first and second explanations. There was a strong relationship between the two values ($\chi^2(9, 313) = 266.16$, $p < .001$, $\varphi c = .532$), suggesting that children tended to give two consistently scored explanations for both dolls.

Table 3.1: Means for each condition for each explanation type

Explanation	Physical				Social			
	Condition	Age	Mean	SD	Condition	Age	Mean	SD
External	Toy	4	1.04	0.91	Toy	4	1.08	0.88
		6	0.75	0.90		6	0.79	0.83
	Control	4	0.88	0.87	Control	4	0.67	0.84
		6	0.72	0.73		6	0.38	0.61
	Doll	4	0.21	0.41	Doll	4	0.29	0.55
		6	0.25	0.53		6	0.22	0.42
Internal	Toy	4	0.75	0.90	Toy	4	0.58	0.78
		6	0.96	0.95		6	1.04	0.86
	Control	4	0.81	0.86	Control	4	1.07	0.87
		6	0.44	0.62		6	1.53	0.62
	Doll	4	1.79	0.41	Doll	4	1.54	0.72
		6	1.21	0.83		6	1.74	0.45
Interaction	Toy	4	0.08	0.28	Toy	4	0.29	0.69
		6	0.17	0.48		6	0.17	0.56
	Control	4	0.22	0.55	Control	4	0.20	0.41
		6	0.66	0.79		6	0.09	0.30
	Doll	4	0.00	0.00	Doll	4	0.13	0.45
		6	0.50	0.78		6	0.04	0.21
Other	Toy	4	0.08	0.41	Toy	4	0.00	0.00
		6	0.13	0.45		6	0.00	0.00
	Control	4	0.09	0.39	Control	4	0.03	0.18
		6	0.19	0.54		6	0.00	0.00
	Doll	4	0.00	0.00	Doll	4	0.00	0.00
		6	0.04	0.20		6	0.00	0.00

3.3.2 MANOVA results

We conducted an omnibus Domain (Psychological, Physical) x Condition (Doll, Toy, Control) x Age (4, 6) x Consistency (Deterministic, Probabilistic) MANOVA on the four scores for internal, external, interaction, and other explanations. There was a main effect of Domain ($F(4, 290) = 4.343$, $p < .01$), Condition ($F(4, 290) = 7.246$, $p < .001$), and Age ($F(4, 290) = 2.812$, $p < .05$), but not of Consistency ($p > .93$). There were two significant interactions: one between Domain and Condition ($F(8, 582) = 2.403$, $p < .05$), and the other between Domain and Age ($F(4, 290) = 5.38$, $p < .001$). Because Consistency had no main or interaction effect across types and no main effect for each individual explanation type, it was excluded from subsequent analyses. [The only significant effect of Consistency was as an interaction with Domain for interaction explanations, $p = .02$.] For the proportions of explanations of each type, see Figure 3.1.

We then did a series of tests contrasting four- and six-year-olds. Across conditions and domains, four-year-olds gave significantly more external explanations than six-year-olds ($F(1, 315) = 4.534$, $p < .05$) and six-year-olds gave more interaction explanations than four-year-olds ($F(1, 315) = 4.072$, $p < .05$). Then we looked at whether children's explanations differed across domains. As predicted, four-year-olds' explanations did not differ across domains ($p > .2$).

Six-year-olds gave significantly more internal explanations about people than physical objects ($F(1, 153) = 23.518$, $p < .002$). Thus four-year-olds and six-year-olds had different responses to changes in Domain (Domain x Age $F(1, 313) = 11.697$, $p < .01$). Whereas the previous study showed that six-year-olds overall had a greater propensity to give internal explanations than the four-year-olds when the dolls were treated as people ($F(1, 155) = 8.898$, $p < .01$), this effect was reversed when the dolls were treated as objects. In this case (the Physical domain), four-year-olds gave more internal responses than the six-year-olds ($F(2, 154) = 3.352$, $p < .05$). (For a graph of these results, please see Figure 3.2.)

Conversely, six-year-olds gave significantly more interaction explanations about objects than people $F(1, 153) = 13.879$, $p < .01$), where they also differed from four-year-olds ($F(1, 315) = 14.444$, $p < .001$). There was no Domain effect on external explanations.

Next, we looked at whether children gave different types of explanations for each covariation pattern, and looked at effects of Condition on each explanation type. Both four- ($F(2, 152) = 21.201$, $p < .001$) and six- ($F(2, 153) = 7.186$, $p < .01$)-year-olds were more likely to give internal explanations in the Doll condition than the other two. Four-year-olds gave more external explanations in the Toy condition than in the Doll or Control Condition ($F(2, 152) = 13.515$, $p < .001$), as did the six-year-olds ($F(2, 156) = 7.208$, $p < .01$). Interaction explanations, which implicate the doll and toy equally, did not differ by covariation pattern.

Within the Doll condition, six-year-olds gave significantly more internal explanations when the dolls were treated as people and not as objects ($F(1, 45) = 7.302$, $p < .05$), even while the data equally supports internal explanations across domains.

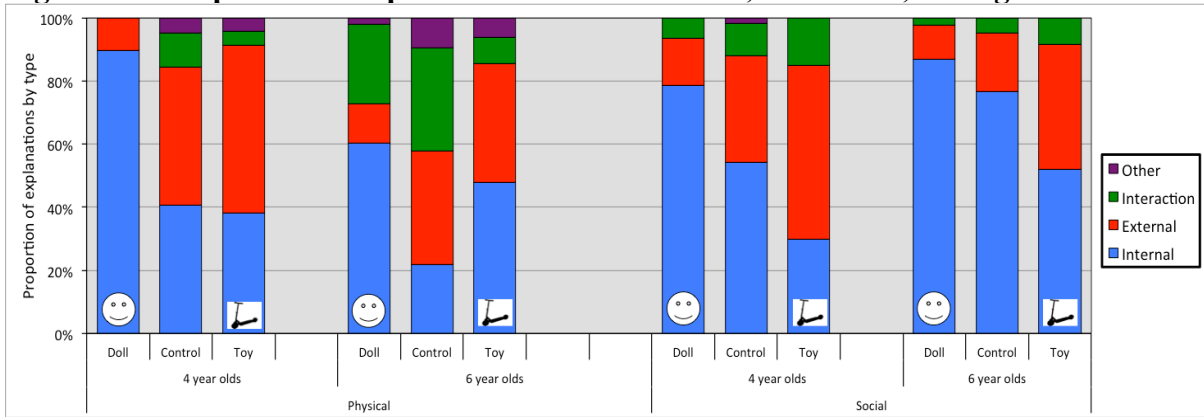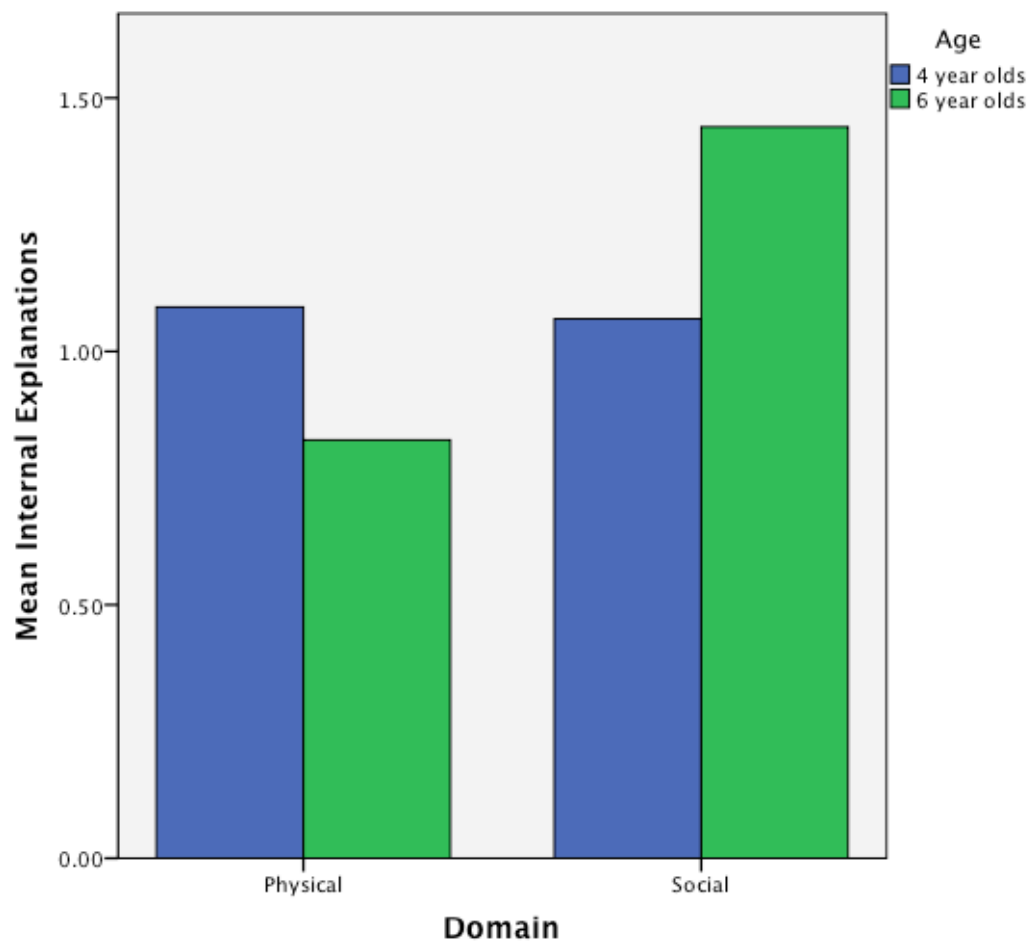**Figure 3.1: Proportion of explanations across domains, conditions, and ages**

Figure 3.2: Domain x Age interaction for internal explanations

3.3.3   Congruency

We also analyzed the explanations in terms of congruent and incongruent answers. A *congruent* response could be construed as following the covariation pattern of the presented data. An *incongruent* response would not follow the implications of the data.

Internal explanation: congruent for Doll condition, incongruent for the Toy condition

External explanation: incongruent for Doll condition, congruent for the Toy condition

We scored having two internal explanations in the Doll condition as a congruent answer, and all other explanation combinations as incongruent. We also scored having two external explanations in the Toy condition as congruent, and all other explanation combinations as incongruent. The control conditions were excluded from the analysis since they have a confounded person by situation pattern that implicates neither.

A congruent score was given the value of '1', and an incongruent score was given the value of '0'. Thus, the mean congruent score conveniently corresponded to the proportion of congruent answers. We then compared these scores between all subgroups. The four-year-olds

gave congruent answers more often than the six-year-olds, except in the Psychological domain Doll condition. When we performed a Domain x Condition x Age univariate test on their scores, there was a large effect of Condition on congruent answers, with significantly more congruence in the Doll condition than the Toy condition (F(1, 183) = 21.949, p <.001). This difference in accuracy across conditions is consistent with the large overall proportion of internal explanations. There was an overall effect of age, with four-year-olds having higher congruence scores than the six-year-olds (F(1, 183) = 4.084846, p = .045). There were no other significant main effects or interactions. Within each condition and comparing across ages, four-year-olds were more accurate than six-year-olds in both the Toy condition (p <.01) and the Doll condition (p = .01), but there were no Domain differences by age. Drilling down further with t-tests contrasting four-year-olds vs six-year-olds by individual Domain x Condition cases, there were two specific cases where four-year-olds had significantly higher congruent scores as determined by t-tests: in the Psychological Toy case (p <.05) and Physical Doll case (p <.01). (For a graph of the congruent data, please see Figure 3.3.)

A "correct" prediction response extended the pattern of covariation information previously viewed. In the Doll condition, this would be predicting that the same doll who played/stuck before would play/stick again on a new toy. If Sally stuck to the scooter and skateboard and Josie didn't stick to either, the child would predict that Sally would be more likely to stick to the newly introduced springboard. In the Toy condition, a correct prediction was that a new doll would behave the same way on each of the toys that the original dolls did. If Josie and Sally both played on the trampoline and didn't play on the diving board, then the child would predict that Mary would also play on the trampoline and not on the diving board. Children in the Control condition did not receive a prediction question since they did not have a meaningful pattern to extend.

For the Physical domain, we asked children a single forced-choice about which doll would stick on the new toy or which toy a new doll would stick to. 71 out of 96 children answered the single prediction question correctly in the Physical domain, a proportion quite likely not due to chance of .5 (z = 4.695, p <.0001) (see Figure 3.4). In the Psychological domain, we asked two independent forced-choice questions. A similar proportion of children predicted a consistent pattern of behavior in the Psychological domain, 46 out of 97 children (z = 4.874, p <.001), although calculated in a different way since there were two prediction questions and chance was .25. (See Figure 3.5.)

A "correct" prediction response extended the pattern of covariation information previously viewed. In the Doll condition, this would be predicting that the same dolls who played/stuck before would play/stick again on a new toy. If Sally stuck to the scooter and skateboard and Josie didn't stick to either, the child would predict that Sally would be more likely to stick to the springboard. In the Toy condition, this was predicting that a new doll would behave the same way on each of the toys that the original dolls did. If Josie and Sally both played on the trampoline and didn't play on the diving board, then the child would predict that Mary would also play on the trampoline and not on the diving board. Children in the Control condition did not receive a prediction question since they did not have a meaningful pattern to extend.

For the Physical domain, we asked children a single forced-choice about which doll would stick on the new toy or which toy a new doll would stick to. 71 out of 96 children

answered the single prediction question correctly in the Physical domain, a proportion quite likely not due to chance of .5 ($z = 4.695$, $p < .0001$). In the Psychological domain, we asked two independent forced-choice questions. A similar proportion of children predicted a consistent pattern of behavior in the Psychological domain, 46 out of 97 children ($z = 4.874$, $p < .001$), although calculated in a different way since there were two prediction questions and chance was .25.

Finally, we compared the congruence scores to the predictions. The congruence score was a significant predictor of the prediction score ($F(1, 187) = 6.153$, $p < .05$). Thus, children who answered in a congruent manner to the data were more likely to also extend this pattern to predictions about future behavior.

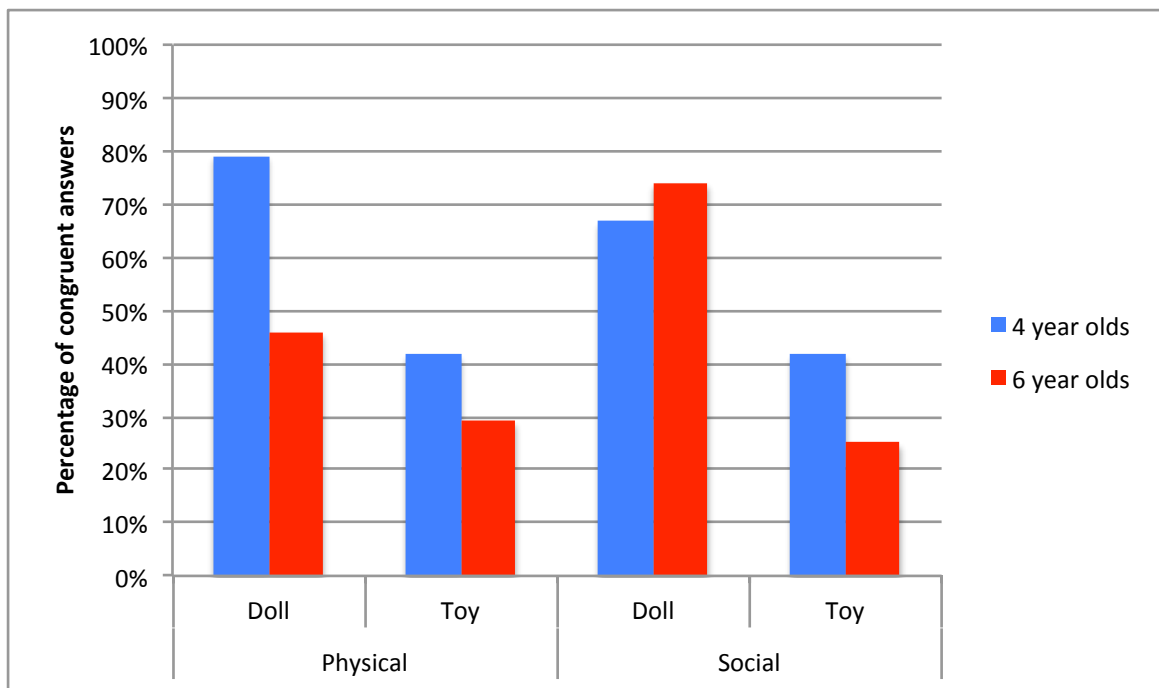**Figure 3.3: Percentage of congruent explanations across ages, conditions, and domains.**

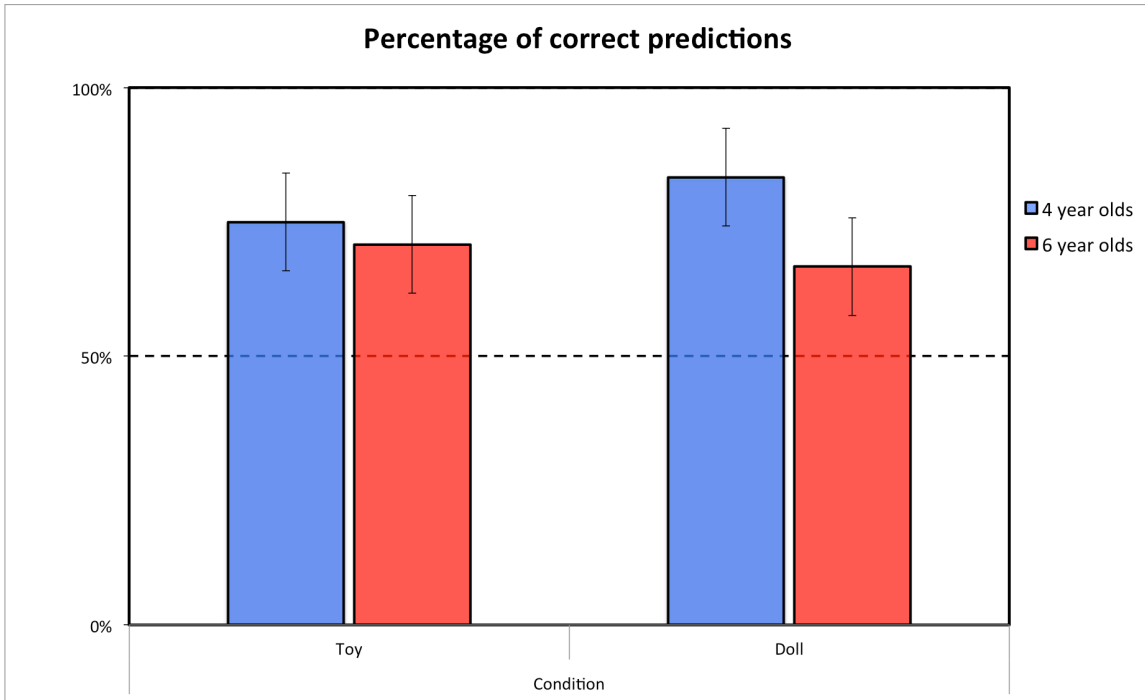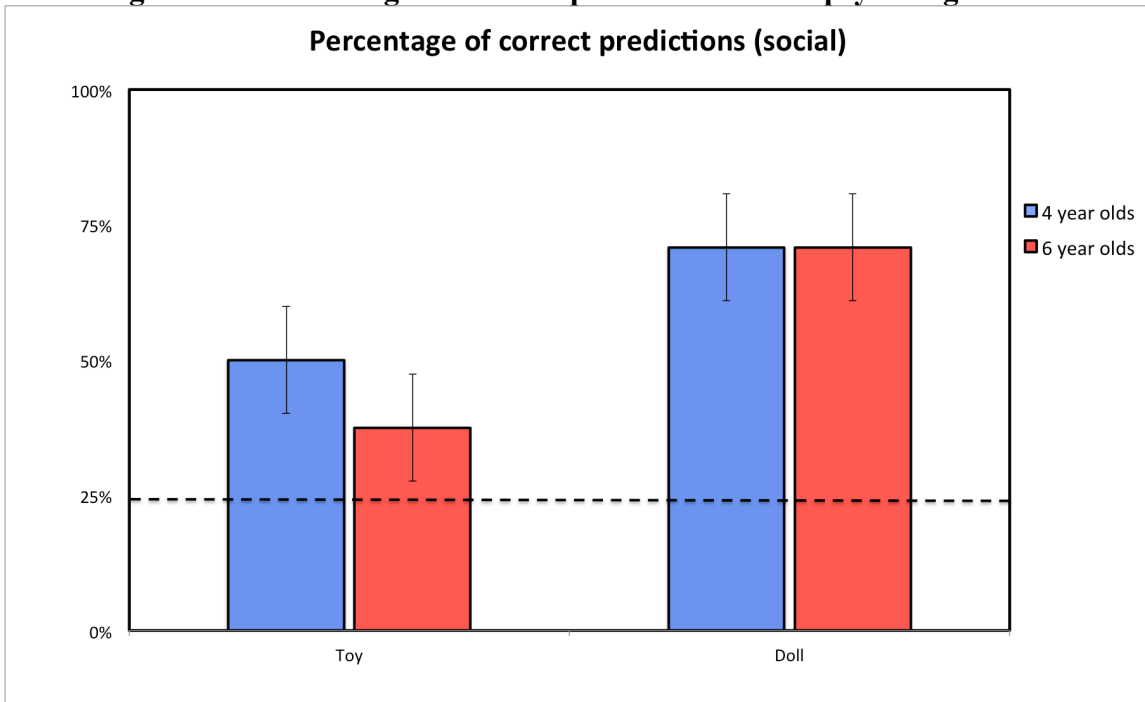**Figure 3.4: Percentage of correct predictions in the physical domain.**

Figure 3.5: Percentage of correct predictions in the psychological/social domain.

3.4   Discussion

The primary factor affecting children's explanation for events was the presented pattern of person x situation (or doll x toy) data; that is, children tended to give answers congruent with the evidence. When actions covaried by Doll in the Doll condition and the two dolls acted differently from each other, children were more likely to give internal explanations. When actions covaried by Toy in the Toy condition and the dolls always worked with one toy and did not work with the other, children were more likely to give external explanations. Children across conditions predicted the same level of consistency for future behavior. They were able to do this even with probabilistic data; that is, children tolerated an amount of noise in the covariation pattern and did not modify their explanations accordingly.

Four-year-olds did not differentiate between the psychological and physical domains for explaining the same covariation pattern. Six-year-olds gave more interaction explanations than the four-year-olds, mostly in the Physical domain and in both control conditions, where the cause is ambiguous. In the absence of informative covariation information, six-year-olds draw on rich, domain-specific prior beliefs for evaluating potential psychological and physical causes. Six-year-olds are both less sensitive to the data than four-year-olds and more discriminating between the two domains.

Importantly, the age pattern of internal explanations differed for each domain. Six-year-olds had displayed an internal bias in the Psychological case that the four-year-olds did not. However, four-year-olds were significantly more likely to give internal explanations than six-year-olds for physical causes. There was no evidence of the internal bias carrying over from the psychological to the physical domain. These results provide evidence that children can both evaluate statistical data and take domain-specific prior knowledge to interpret that evidence.

While Seiver et al (2012) found that six-year-olds had an overall bias towards internal explanations when explaining other people's behavior, the present study suggests that this bias does not extend across domains. With only a change in causal mechanism, six-year-olds appear to track the same covariation pattern differently for the dolls and toys when they are presented as objects than when they are treated as representatives of real people.

Six-year-olds are using strong, domain-specific priors in both the physical and psychological conditions. These prior beliefs implicate internal causes in the psychological domain. For the physical domain, the priors originate from the scientific properties of magnetism. A single magnet can act as a cause on a metal object, or two magnets can stick together. In either case, both the doll and the toy would require a certain property to stick together and thus support interaction explanations.

We also have evidence from the qualitative content of their explanations that the four-year-olds were not as knowledgeable as six-year-olds about magnets. These younger children tended to describe a doll's stickiness related to surface adhesion, as more of a sticky gum or glue ("Cause his feet are sticky"; "cause there is tape on the bouncy board"; "because the skateboard has glue on it") and not as magnets. Under this interpretation, either the doll or the toy could be sticky. Younger children perhaps viewed a single object as having enough adhesiveness that both objects did not need to be sticky.

These data also still do not completely resolve the question of the relationship between the social and physical domains of causal reasoning. One initial hypothesis was that there would be a "bleeding over" effect of psychological domain biases into thinking about the human-like toys. Since we were showing the same visual information and merely varying the domain, we can consider where the line actually lies between psychological and physical causation. Perhaps there is a state between the two that the dolls can inhabit. We might have found domain differences because the dolls in the physical domain do not serve as representations or symbols, but rather as actual physical causes. While most of the evidence in the physical domain is immediately observable (and indeed, children did try to look under the toys and the dolls' feet for magnets), in the psychological domain children are filling in the blanks with their imagination.

The differences in performance across ages exemplifies the efficiency/accuracy trade-off of relying on heuristics and prior beliefs. Six-year-olds have more fleshed out, thoughtful, articulate reasons to explain people's behavior, as well as more sophisticated formal scientific understanding of how magnets work. However, in evaluating an individual's behavior in a situation, this makes it more likely that they will fall back on previous assumptions and over-generalizations than analyze the immediate evidence. Thus we would expect younger children to make more appropriate but also more superficial judgments in the here and now. But these younger children have not yet developed a rich set of theories to think more deeply about the causes of people's actions.

Another interesting direction for future work would be to repeat this study with blickets and blicket detectors: one further step removed from the anthropomorphism of the physical events with dolls. Incidentally, when we attempted to pilot that very study with four-year-olds, they were at chance for implicating the block (person) or the machine (situation). Thus among the many different factors at work, there may also be an interest factor where children are more engaged when presented with human-like actions.

Future work in the psychological domain could examine whether generic situations or specific objects are necessary for tracking behavioral patterns; that is, many situations or environments are more abstract than a specific toy. It also remains to be seen what the ramifications are of more closely following the behavioral data or not. Perhaps this paradigm could be employed in other related studies, looking at heavily-valenced traits or in-group and out-group membership. It is quite possible that children could track qualities of people (be it shirt color or hair color) that would affect their judgment of straightforward correlational data. Learning to track data more carefully could lead to an overturning of such unhelpful priors as stereotypes.

Perhaps older children would show evidence of domain carry-over for an internal bias as posited by Peng and Knowles (2003). Six-year-olds may not have had enough exposure to a culturally biased way of thinking. In most of the studies that provide evidence for attributional biases affecting general causal inference, the study participants are adults or older children than in our sample. Future studies applying this study's paradigm to older ages, as well as to different modalities, would help our understanding of the developmental trajectory of the relationship between culturally-mediated attributional biases and physical causal reasoning.

These studies provide evidence for domain-general statistical reasoning, which is influenced by domain-specific prior beliefs. This fits nicely with the causal Bayes nets account of learning, and suggests several potential mechanisms for theory/stereotype/bias transmission. One possibility is that six-year-olds have acquired more behavioral data overall that has been filtered through an initial, weaker bias, thereby strengthening it. Another possibility is that six-year-olds have acquired their biases via independently strengthened prior beliefs, such as cultural testimony about the causes of behavior or formal scientific training. Cross-cultural work addressing these two possibilities is presented in the following chapter.

Chapter 4   Cross-cultural differences in children's attributional styles

4.1   Introduction

In the previous chapters, we have considered both age differences in attribution as well as domain differences. However, the origin of the prior beliefs of the four- and six-year-olds and what underlies that developmental change is still unknown. How are six-year-olds forming their beliefs? So first there was children tracking people's behavior in different patterns and we found age differences, where older children have stronger prior beliefs and are less sensitive to the data.

There are multiple ways that four- and six-year-olds differ. For example, six-year-olds have been in the world for a longer period of time and thus have had the chance to observe more person by situation data. But other factors are influencing their judgments. One of those is the general cultural environment in which the children are raised. They are exposed to and seek out additional adult explanations for others' behavior (as young children are often known to pester their parents with the question "why?"). This is part of the general cultural environment that children grow up in, a sea of underlying attitudes that are shared by many of the people around them. Some of these attitudes are the tendency to interpret and explain people's behavior in a certain way.

The fundamental attribution error was first described in Jones & Harris' (1967) landmark study where adults who read an essay for or against Castro and were told that the opinion slant was determined by the flip of a coin, still thought that the pro-Castro essay writer favored Castro and the anti-Castro writer was against Castro. In this specific case where there is a situational explanation for the opinion slant and technically no relevant dispositional information, it is more correct to categorize it as an "error".

Why was FAE renamed the correspondence bias? It is not a universal error as cleanly presented in the (Jones & Harris, 1967) paradigm. Sometimes, it makes sense for us to infer traits or make dispositional inferences from behavioral evidence. This is especially true when there are no circumstances presented as a plausible alternative explanation, such as when someone is behaving in an anti-social manner. Sometimes behaviors do correspond to traits. However, adults show a tendency to infer that a behavior is caused by a trait more than warranted by the evidence. To put it in terms of causal inference, a tendency to overestimate the probability of a potential cause is merely a bias instead of a factual error. Adults have strong prior beliefs about the causes of people's behavior that interact with their processing of more immediate and relevant data.

It is important now to take a step back from the previous studies to point out that these children studied thus far are all from the United States. Thus they are specifically a product of American culture. Might we see cultural differences in other countries where adults engage in a different attributional style? It could be a universal developmental trend. However, based on findings in cultural psychology, it is reasonable to believe that this may be due to unique features of American cultural attributional styles that are not universal. We have said that the evidence pointed to an early emergence of the fundamental attribution error. Thus a relevant cultural comparison would be with one that the error does not appear to the same extent.

One of the countries that is often contrasted in cultural psychology with the United States is China. These countries are both quite vast and diverse, and there are many within-cultural differences in social cognition. However, there tend to be some overarching commonalities within these cultures as well, especially in comparably metropolitan areas. Some of the study findings with mainland Chinese adults also have been found in other countries like Korea and Japan, so I will be citing some broader studies that incorporate adults from other areas of East Asia, including analysis of the area of East Asia as a comparison cause itself for the United States and European/Western set of beliefs.[1]

The correspondence bias, though highly prevalent in the United States, is not culturally specific, but a universal phenomenon that occurs across cultures. As was mentioned earlier, one of the attributional biases more prevalent in American than Chinese culture is the correspondence bias, where adults tend to attribute people's behavior to stable and underlying traits with little supporting evidence. However, although Western adults have correspondence bias more than East Asian adults, both groups reason in this biased way (Krull et al., 1999). Why do Westerners have a stronger bias?

It appears that correspondence bias is a kind of default that is mitigated by other evidence. Different situations bring it out more than others, and some external circumstances are only mitigating for members of certain cultures. East Asian adults are better at attenuating the fundamental attribution error with situational evidence.

Dispositionism is also universal, but the underlying assumptions about the reliability and consistency of dispositions vary across cultures. Especially in the case of traits, one's belief that there is a trait of intelligence has a number of potential theories about how that intelligence came to be. For example, one might believe that an IQ score is determined by your genetics. However, you could also believe that intelligence is acquired through fortunate SES or upbringing, or hard work and dedication. In essence, this difference between the United States and East Asia corresponds to Dweck's (cf. Dweck, Hong, & Chiu, 1993) entity and incremental theories of traits (Choi, Nisbett, & Norenzayan, 1999) and to the cultural distinction at hand.

So then American adults' bias cannot be explained by dispositionism. As we are considering Kelley's person by situation construct, perhaps adults in the United States are not sensitive to information about the situation. It is also possible that East Asians have a situational bias. Both could explain why Americans cannot mitigate the fundamental attribution error as well as East Asians. If the former is true, the ability to use situational evidence may be entire learned by enculturation, so the United States adults simply do not understand how to apply this evidence to their attributions. If the latter is true, an alternative possibility is that the East Asian situational bias is part of a higher order theory via

---

[1] It has been pointed out that the East/West distinction is rooted in fairly biased American-centric thinking about how people behave in collectivist and individualist cultures, oversimplifies intra-cultural distinctions, and also leaves out many other areas of the world with their own distinct and comparable attributional styles. Although I won't be directly addressing this very real and pervasive problem in psychology research, see Yong (2009) for more information on this issue.

enculturation that doesn't actually constrain their attribution powers per se but rather tends to constrain the information that adults consider in making their judgments.

Some studies show that East Asian adults have a situational bias and that American adults can use situational information. When Americans are put under cognitive load and presented with a an attribution task that emphasizes the situation, they use situation information *better* than if they have access to their full cognitive faculties (Lieberman, 2005). When East Asians are put under cognitive load, they are more likely to make dispositional attributions (and to discount dispositional information if they are not), demonstrating that they too have acquired an attributional bias. (Lieberman, 2005) This suggests that higher-order, well-developed theories about behavior are influencing adults' attributions across cultures but the actual theories are specific to certain cultures.

Less known is the effect of culture on children's causal reasoning. First, they are younger and less entrenched in a cultural bias. According to some studies (e.g., Miller, 1984) young children do not have a bias when reasoning about people's behavior. However, Seiver, Gopnik and Goodman (2012) showed that there were glimpses as young as six years of age in a preference for explaining behavior in terms of the person, even with behavioral evidence implicating the situation as a cause. This suggests that young children in the United States might perceive physical events through the lens of an emerging correspondence bias.

In children, some work has looked at how theory of mind develops cross culturally. Theory of mind is the general understanding that other people have beliefs and desires and general mental states that are independent of their own beliefs, and they can hold incorrect beliefs that differ from reality.

Children across cultures follow the same general trajectory of increasing theory of mind understanding (Shahaeian, Peterson, Slaughter, & Wellman, 2011). However, there is some difference between cultures for the average age of attainment for specific tasks (Liu, Wellman, Tardif, & Sabbagh, 2008). For children in the United States, they understand opinion diversity (that two people can hold different opinions simultaneously) before they understand knowledge access (that someone else may not know what they know) (Wellman, Fang, Liu, Zhu, & Liu, 2006). For children in both China and Japan, they attained the tasks in the reverse order. This is in line with each of the cultures: opinion diversity matters more in a culture where people are attending to dispositions, and less so in a culture where there is an expectation of greater harmony within the group. For knowledge access, if East Asians are better at perspective taking as we already know they are in adulthood (Wu & Keysar, 2007), than it makes sense that they would pass this false belief test before Americans would. This provides evidence that overarching cultural theories are already impacting how young children learn about other people's minds.

In Joan Miller's (1984) work, children are diverging at an older age in the United States and South Asia in attributional styles. Thus we have reason to believe that the children are not growing up to favor internal explanations for people's behavior, and that six-years-olds would also be too young for spontaneous trait words. So we can find that rich middle ground in the United States, where children are familiar with and understand trait words but aren't using them (Heyman, 2009; Liu et al., 2007). We can glean the beginnings of biased thinking in any direction by examining how closely explanations conform to the given person by situation covariation information. The Chinese six-year-olds potentially show evidence of

developing other prior beliefs that may coalesce into stronger biases later in childhood. We still need to test whether the internal bias from Chapter 2 is symptomatic of an emerging correspondence bias.

There are multiple possible outcomes when comparing across cultures. First, because of less internal bias, Chinese six-year-olds may look more like the American four-year-olds. But it is also possible that they may be picking up biases unique to mainland Chinese culture. So if the Chinese and US four-year-olds look similar, and the United States and Chinese six-year-olds diverge from the four-year-olds, we would have cross-cultural evidence of the development of stronger, culturally specific prior beliefs about people's behavior. If the United States and China six-year-olds perform similarly, that would suggest that the internal bias is not related to the development of fundamental attribution error. Person x situation covariation information would then seem to be the strongest indicator of this type of bias. To this end, we repeated Seiver, Gopnik, and Goodman (2012) in adapted form for mainland China.

4.2   Method

4.2.1   Participants

Across conditions there were 91 four-year-olds (M = 4.33 years, range = 3.92 − 5.00 years) and 92 six-year-olds (M = 6.14 years, range = 5.83 − 6.96 years). Children were recruited from area preschools in the city of Beijing, China, as well nearby suburbs. The experimenters were all female and native Mandarin speakers. They were undergraduate students from Tsinghua University and University of California Berkeley.

The procedure was virtually identical to Seiver, Gopnik, & Goodman (2012). We slightly modified the protocol to adjust for cultural differences between the United States and China. First, several graduate students at Tsinghua University, whose native language was Mandarin and who were highly proficient in English, translated the procedures from English to Chinese. A separate set of graduate students back-translated the procedures from Chinese to English, and any discrepancies between the original and back-translated English versions were resolved in the Mandarin translation through a group discussion.

4.2.2   Test Design

Person, Situation, and Control Conditions

In each experimental condition, participants observed a total of eight engaging actions and eight backing away actions. However, the distribution of those actions either covaried with the situation or with the individual (person) in a between-subjects design. In the person condition, for example, Huanhuan would consistently play on the bicycle and the trolley, while Beibei would consistently back away from both activities. In the parallel situation condition, both Beibei and Huanhuan would consistently play on the bicycle, but they would both consistently back away from the trolley. In the control condition, children did not get a full set of information as in the experimental conditions, to get a baseline for preferences of explanation type. In the corresponding control condition to the previous examples, a child would see Huanhuan play on the bicycle four times and then Beibei back away from trolley four times. This control condition contains the eight events that are identical between the

person and situation condition. Importantly, every aspect of the procedure was held constant between these conditions except for the actual covariation pattern.

Deterministic and Probabilistic Frequency

We also varied the consistency of the actions, to mimic real life where a person might not be 100% consistent in their actions, or not everyone will always do the same thing in a given situation. In the deterministic conditions, the dolls either engaged in or backed away from the activity consistently on all four trials. In the probabilistic conditions, the dolls either engaged in the activity three out of four times or backed away three out of four times. The anomalous evidence occurred on the third approach.

4.2.3 Materials

We adapted the dolls and activities to be more culturally appropriate for Chinese children, using toys from local Beijing markets. We named the three dolls Beibei, Huanhuan, and Nini, familiar child-like names in Mandarin. The activities that the dolls played on were a wagon, a bicycle, and a springy dog toy, much like a stationary ride at an amusement park, that the dolls could sit on.

All children who participated were in a separate designated classroom, run by a female experimenter, and randomly assigned to each experimental condition. Participants observed a vignette in which two characters (the dolls Beibei and Huanhuan) made a series of approaches to two activities (chosen from among the trolley, bicycle and spring dog). The first doll would approach activity A four times, followed by the second doll, who would also approach that activity four times. Then the first doll would approach activity B four times, followed by the second doll. On each trial the doll would either engage in the activity (dive in the pool, jump on the trolley or ride the bike), or else would back away. Doll order and activity order were counterbalanced across participants.

For example, in the deterministic person condition, children might see Huanhuan jump on the trolley four times, then see Beibei approach the trolley and back away four times, then see Huanhuan ride on the bicycle four times and then see Beibei approach the bicycle and back away four times. The experimenter narrated throughout on each trial as each doll either engaged in the activity or backed away (e.g. "Look! Huanhuan's playing on the trolley. She's not scared." in the deterministic condition or "Look! Huanhuan's playing on the trolley" in the probabilistic condition).

Explanation questions

At the end of the vignette, the experimenter asked two open-ended explanation questions, one for each doll's last action on the second activity (e.g., "Why did Huanhuan play on the bicycle?" and "Why didn't Beibei play on the bicycle?"). Children first answered for the doll most recently viewed, and then answered for the second doll. Children were free to explain the behavior as they wished. If the participant refused to answer, or gave an irrelevant answer, the experimenter would follow up with a forced-choice question contrasting a person and situation attribution (e.g., "Why did Huanhuan play on the bicycle? Is it because she's the kind of person who does brave things, or because the bicycle is safe to play on?" or "Why didn't Beibei play on the bicycle? Is it because she's the kind of person who gets scared, or because the bicycle is dangerous to play on?"). Children sometimes

responded to the explanation questions by simply saying, "Because she wanted to" or "because she didn't want to", especially in the probabilistic condition. The experimenter followed these responses with the question "Why did she want to?" The follow-up answer was used for coding the explanation type. A scoreable response was needed before moving on to the next question.

Prediction questions

In the person and situation conditions, the experimenter then asked two prediction questions about a new future event. We tested whether children in those conditions would generalize from the pattern that they observed. In the person condition, the participants were asked to predict what each doll would do in a new situation ("Now let's pretend that Beibei and Huanhuan go over to the spring dog. Do you think Huanhuan will play on it? <child answers> Do you think Beibei will play on it?"). In the situation condition, children were asked to predict what a new doll would do in each of the earlier situations ("Now let's pretend that Beibei and Huanhuan have a friend named Nini. Do you think she'll play on the trolley? <child answers> Do you think she'll play on the bicycle?"). Children answered either "yes" or "no." For predictions to be scored as correct, children had to answer both questions in the manner consistent with what they had previously viewed. All other patterns of responses were scored as incorrect. Thus if children were responding at chance, they would receive a correct score 25% of the time.

As noted above, if children did not provide a relevant explanation spontaneously, they received the forced-choice question. Both forced-choice and open-ended explanation responses were transcribed, translated, and coded into four mutually exclusive response types. An explanation was coded as an "internal" response if it attributed playing to the doll. This cause could involve physical characteristics of the doll such as age or size, or mental states such as her desires, beliefs, or intentions. An explanation was coded as an "external" response if it referred to an underlying cause that was outside of the doll, including the toy she played on. An explanation was coded as an "interaction" response if it explicitly cited both the doll and the toy, such as "she likes it and it is fun to play on the dog". Any explanation not codeable into those three categories was coded as "Other". Finally, forced choice responses were coded "internal" or "external" based on which of the two options was chosen. All explanation translation and coding was done by native Mandarin speakers from the original audio files. Two coders independently listened to each audio file, to make sure we had reliability and consistency in the translations; there were five coders in total.

4.3   Results

The coding was then compared to the Seiver, Gopnik, & Goodman recoded data with children from the UC Berkeley area. (For details on the original study, see Chapter 2. For details on adapting the data in Chapter 2 to the newer coding system, please see Chapter 3.) The results that follow are from preliminary analyses.

First we ran an overall Country (United States, China) x Condition (Person, Situation, Control) x Age (4, 6) x Frequency (Deterministic, Probabilistic) MANOVA for Internal, External, Interaction, and Other scores.

Internal explanations

The number of internal explanations differed significantly for each Condition ($F(2, 330) = 19.025$, $p <.001$), with the most in the Person condition. Across both countries and all conditions, six-year-olds gave internal explanations more frequently than four-year-olds ($F(1, 330) = 6.813$, $p <.01$); Just at significance was a Country x Age interaction $F(1, 330) = 3.813$, $p = .052$. There was also an interaction between Country and Condition ($F(2, 330) = 4.516$, $p <.05$).

Figure 4.1 Age differences across countries for internal explanations
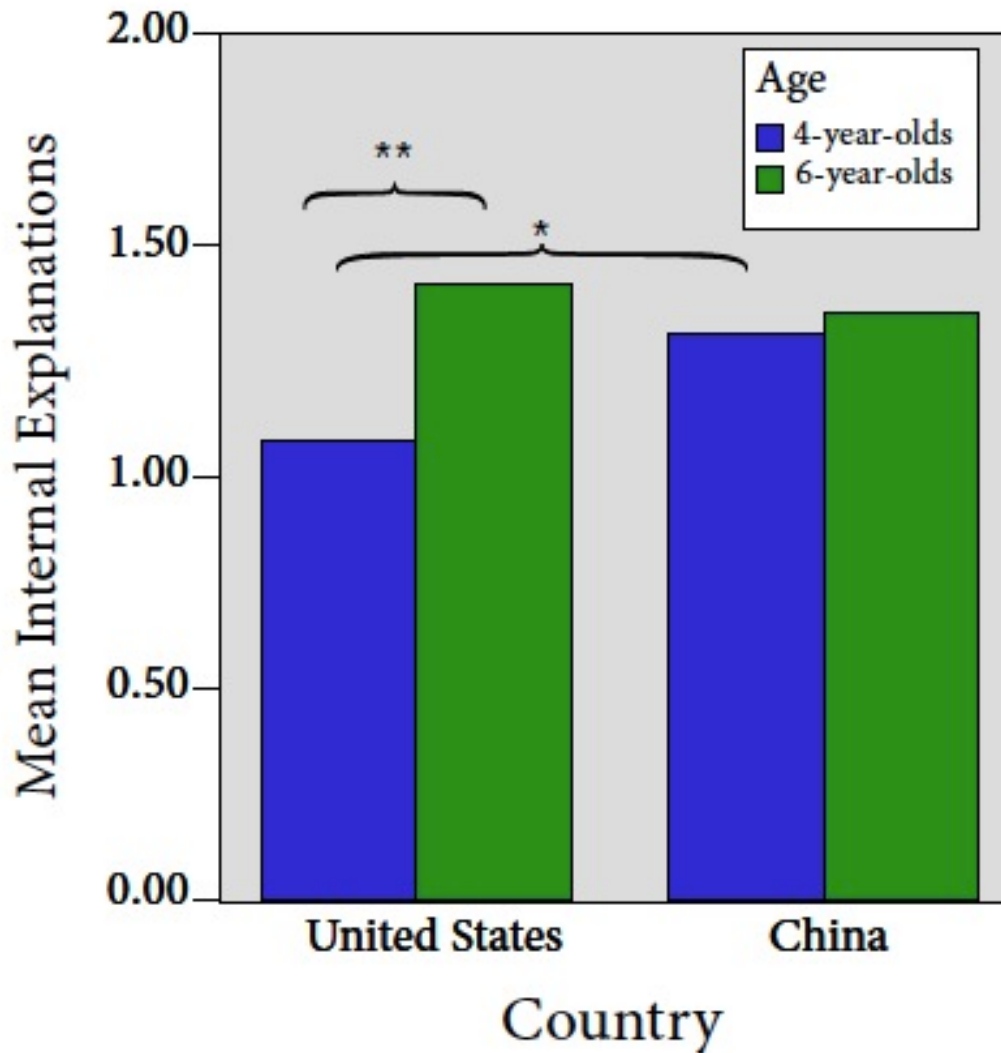
External explanations

The smallest number of external explanations occurred in the Person condition ($F(2, 330) = 14.013$, $p < .001$). The condition with the largest number of explanations was significantly different across the two countries, with the most external explanations in the Situation condition for children in the United States, and significantly more external explanations for the children in China in the control condition ($F(2, 330) = 3.778$, $p < .05$).

Interaction explanations

For Interaction explanations, there were significantly more for the deterministic than for the probabilistic frequency ($F(1, 330) = 6.396$, $p < .05$).There were no other main effects or interactions (i > .5 for all). Since frequency had no other significant effects, both interaction explanations and frequency information were excluded from future analyses.

Other explanations

Because there were so few explanations in the Other category (M = .01 out of a possible 2), it was also excluded from subsequent analyses.

4.3.2 Planned comparisons

For planned comparisons, we focused on internal and external explanations as they corresponded the most closely to person and situation. (In Chapters 2 and 3, it was not significant for any explanation type.) In Chapter 3, interaction explanations were tied closely to physical causation, so it is not surprisingly that we had few in this study.
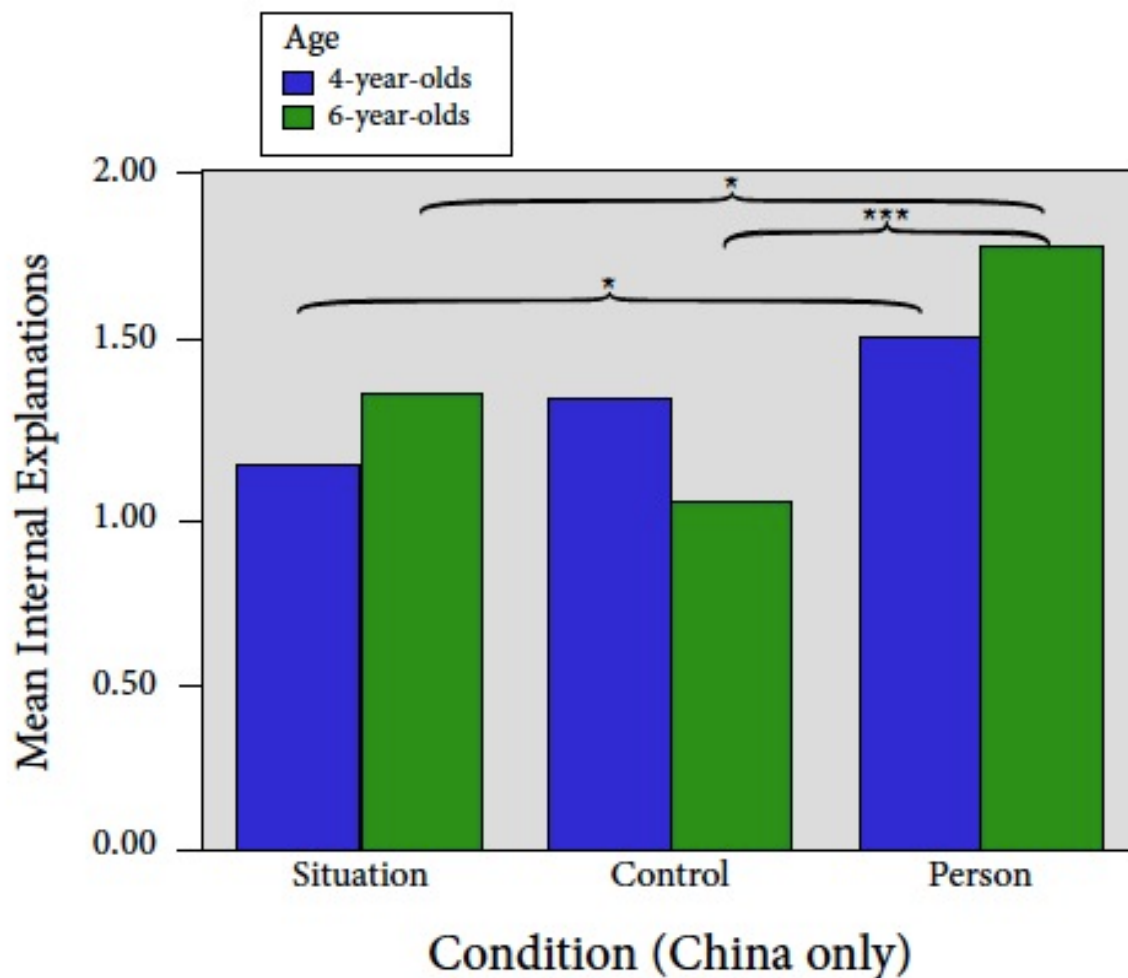
Internal explanations

We ran an additional ANOVA for just internal explanations to detect any other noticeable effects. This category was the largest explanation type across Conditions, accounting for over half of all explanations. Across both countries and age groups, children demonstrated a sensitivity to the covariation pattern. That is, children tended to give more internal explanations in the Person condition, fewer internal explanations in the Control condition, and fewest in the Situation condition ($F(2, 342) = 18.804$, $p < .001$).

Six-year-olds gave more internal explanations than four-year-olds ($F(1, 342) = 6.449$, $p < .05$), though this differed somewhat across countries ($F(1, 342) = 4.606$, $p < .05$). (See Figure 4.1 for a graph of these results.) While the United States six-year-olds gave significantly more internal explanations than the four-year-olds ($t(155) = -2.983$, $p < .01$), the Beijing sample showed no age difference in internal explanations. And while the six-year-olds looked the same across cultures, the China four-year-olds gave significantly more internal explanations than the United States four-year-olds ($t(170) = -2.532$, $p < .05$). There was also a significant Country x Condition interaction ($F(2, 342) = 4.412$, $p < .05$).

We also analyzed the Beijing sample separately. For these children, there was an effect of Condition ($F(2, 191) = 6.181$, $p < .01$) (See Figure 4.2). Both the four- ($t(68) = -2.022$, $p < .05$) and six- ($t(64) = -.2.256$, $p < .05$)-year-olds gave significantly more internal explanations in the person condition than the situation condition. The largest Condition difference for the six-year-olds was where they gave significantly more internal explanations in the Person condition than the Control condition ($t(67) = -3.757$, $p < .001$). While the United States six-year-olds gave consistently more internal explanations than the United States four-year-olds, there was no significance difference within any given condition for the China four- and six-year-olds.

Across conditions, younger children from Beijing followed a consistently rising pattern of internal explanations following the data. However, the six-year-olds had a different shape, where they gave the least internal explanations in the control condition.

Figure 4.2: Condition and age differences within the China sample

External explanations

　　We then examined external explanations separately. Children gave the most external explanations in the Situation condition, and the fewest in the Person condition ($F(2, 342) = 14.101$, $p <.001$). However, this differed across countries (Country x Condition $F(2, 342) = 3.585$, $p <.05$). Both countries were similarly low in the Person condition, but in China there were more external explanations in the Control condition than Situation condition.

　　We then examined children separately at the different ages. Since six-year-olds in the United States displayed an overall bias for internal explanations, we wanted to see if this bias remained consistent across cultures. Six-year-old children responded to the covariation significantly differently across cultures for both internal ($F(2, 176) = 4.750$, $p <.05$) and external ($F(2, 176) = 4.681$, $p <.05$) explanations. Children in the United States showed a steady increase in internal explanations from the data not supporting them (Situation condition), neutral (Control) and supporting (Person). Six-year-olds in China, on the other hand, had a noticeable dip in internal explanations and increase in external explanations for the control condition.

We then examined the four-year-olds' explanations separately. For internal explanations, the Beijing children gave significantly more than United States ones ($F(1, 166)$ = 6.922, p <.01). For both internal ($F(2, 166)$ = 10.298, p <.001) and external ($F(2, 166)$ = 7.043, p <.01) explanations, children followed the general pattern of covariation information.

As in Seiver et al (2012), children received a score of 'correct' for prediction if they predicted that a new doll would do what the first two did (Situation condition) or they predicted the two dolls would continue to act the same way on a new toy (Person condition). (There was no prediction for the Control condition.) Children in either case received two yes/no questions ("Do you think Beibei will play on the bicycle? Do you think Huanhuan will play on the bicycle?"). Thus answering each randomly would be 25% correct.

For deterministic data, there was a greater likelihood of predicting consistency than probabilistic data ($F(1, 210)$ = 4.334, p <.05). This suggests that while children can handle noise in behavioral data, this noise affects their inferences about future events, taking statistics into account. For just the Chinese children, whether they predicted a consistent pattern depended even more strongly on statistical information ($F(1, 126)$ = 11.749, $p <$ .001). The US children predicted more consistency than the children in China ($F(1, 210)$ = 5.246, p <.05).

By far, the strongest effect on consistent predictions was whether it was about a novel person or a novel situation. Children across cultures were more likely to extend the data pattern for a novel situation with the same two characters than a novel person with the same two toys ($F(1, 210)$ = 61.791, $p <$ .001). This was also true for just the Chinese children ($F(2, 126)$ = 20.427, p <.001).

As in Chapter 3 of this volume, we created a score called "congruence" for the two test conditions. For the Person condition, explaining the behavior congruently with the person by situation covariation information would involve giving two internal explanations. For the Situation condition, a congruent answer would consist of two external explanations. This allows us to measure how well the children were following the data by combining information across multiple explanation types.

As with the predictions, children gave significantly more congruent answers for the person condition than in the situation condition ($F(1, 210)$= 62.029, $p <$ .001). This fits with the overall favoritism towards internal explanations.

We then compared the relationship between congruent explanations and predictions. We expected there to be a positive association between the two, as both serve as complementary facets of attribution: locus of cause, and the stability of that cause. There was a strong correlation between the two values (r = .298, p <.001). This correlation significant for all major subgroups, but was noticeably higher for six (.526) than four-year-olds (.210), and higher for China (.455) than the United States (.210) sample. The China six-year-olds were at r = .653, and the China four-year-olds were at r = .278.

4.4   Discussion

Children across cultures from at least as young as four years of age are able to follow person by situation covariation information to guide their attributions for the causes of a person's action in a given situation. Comparing internal and external causal explanations effectively captured the tracking of this statistical information. While there was an overall

preference to explain behavior in terms of the person, there was a developmental shift in the United States children to prefer internal explanations, there was a stronger baseline in the control condition for Chinese six-year-olds to prefer external explanations. Overall, the most significant factor for explanation type, congruence, and prediction is condition, where children are significantly more likely to extend the patterns for people than for objects. So in all cases, the covariation data is having the strongest effect on their attributions.

Given the high overall prevalence of internal explanations, it is safe to assume that across cultures, children find that it is reasonable to attribute cause to the agent of an action. However, this judgment is still sensitive to person by situation covariation data. Children can temper their pre-existing beliefs with immediate covariation information.

Whereas in the United States there was a developmental shift towards increased internal explanations, there was no significance difference between the Chinese four-year-olds and six-year-olds on internal explanations.

Intriguingly, interaction explanations are not related to culture or age for talking about people's behavior, but rather depend on the strength of the relationship between the two. Given that these two variables have a nearly exclusive relationship in the context of this study, this suggests that interaction explanations and frequency information may not be related to person by situation information. Frequency is also related to the stability of inferences about future behavior, as evidenced in the children's higher consistent predictions for deterministic than for probabilistic data.

United States six-year-olds maintained a high level of internal explanations across conditions, including in the control condition. The China six-year-olds' comparatively higher preference for external explanations suggests that they had different prior beliefs than the United States six-year-olds. The shift in internal explanations from the control to the person condition was very dramatic for Chinese six-year-olds, which suggests that an expectation for internal explanations did not form the basis of their judgments and was only activated in the presence of supporting data.

Because the China six-year-olds do not follow the same pattern as the United States six-year-olds, this also suggests that they are operating from a different set of prior beliefs where they are less focused on internal reasons to explain a person's behavior. And indeed, there is evidence that older children have more a consistent theory about behavior than younger children, including within the Chinese sample. Six-year-olds were more consistent than four-year-olds with their explanations and predictions. While we originally had data in the United States where four-year-olds were following the data better than six-year-olds, the six-year-olds appear to be more internally consistent with their theory (whether it matches or doesn't match the data). This suggests a greater degree of insight and richer theory about behavior.

Together with the Peng data that shows the same notion of agency but different descriptions of newtonian mechanics, it is unclear whether the overall preference for internal explanations is a bias and how much it is a reflection of assigning responsibility to an agent rather than a patient. As the United States children do, the children in China overwhelmingly believe that people are responsible for the causes of their own actions. However, the degree to which they believe a person is responsible changes with supporting, neutral, or contradicting evidence.

If children are more willing to extend predictions to people than to objects, this also seems to indicate a domain difference. From Chapter 2, we know that the children have different prior beliefs about each domain. And with contrasting the dolls and the toys, although we framed this experiment in terms of psychological causation, it is still contrasting the physical (toy) domain with the person (doll) domain. These different kinds of prior beliefs appear to hold across cultures with different attributional styles. This is yet another way to see how prior beliefs affect judgments. To the children, a new person is more of a 'wild card' than two people in a new situation.

At a younger than previously thought possible, we have demonstrated several important findings. One is that we can use statistical reasoning to see how attributional biases such as the correspondence bias are formed instead of recording explicit trait language. Two, we have a plausible mechanism for how this might form, including the influence of culture on prior beliefs. Even by the age of six, children in the United States have a preference to explain people's behavior in terms of inherent properties, thoughts, or characteristics. Children in mainland China, on the other hand, have a stronger baseline preference than the United States six-year-olds for explanations about the situation or environment. Although there is a general preference for explaining the behavior with internal explanations that can be partially explained by the causal asymmetry of a person and an artifact, this cannot account for the differences across ages, cultures, and statistical covariation information. By using a more implicit measure of attributional bias (or rather, attention to person by situation covariation information), we have evidence that even young children are already absorbing new information in a way that has been shaped by their environment and their culture.

Chapter 5 Conclusion

These three studies together provide a rich picture of the development of social causal inference. In the first study, we found that four-year-olds follow behavioral data more closely than six-year-olds. While older children had a bias to explain behavior in terms of the person, they also gave richer and more detailed explanations of behavior. When we then extended the procedure into the physical domain, four-year-olds were still following the data more closely than six-year-olds. However, the six-year-olds were operating from a different set of domain specific prior beliefs. They used their understanding of how magnets work to implicate both the doll and the toy playing a causal role when they would stick together or bounce apart. And finally, we found that the internal bias observed in American children living in the Bay area did not extend to children in mainland China, where adults are less likely to make the correspondence bias than in the United States. However, Chinese six-year-olds were more consistent in how they explained and predicted behavior than the four-year-olds, but had fewer congruent answers that neatly lined up with the different conditions.

There are several directions to take this research for further study of attribution from a causal inference perspective. First, we could make stronger developmental claims by extending to older children. The procedure would need to be modified to be developmentally appropriate, perhaps with a video display or less direct narration of the actions. If eight and ten-year-olds were tested, we could get a sense of whether the internal bias persists at older ages and also examine their speech for more overt trait language.

Extending these studies to adults in a culture either high or low in the correspondence bias would provide further of evidence of whether this internal bias is indicative of an emerging trait bias. Again, the procedure would need to be modified. Contrasting adults' performance on the standard task presented in Chapter 2 who were from a corresponding bias prone and not-prone culture.

For Chapter 3, the magnet paradigm could be expanded to other forms of causality. Lewin (1934) argued that certain aspects of Newtonian physics, such as flotation, tap into other attributional biases. He said that adults would be more likely to think of a wood object floating as possessing a property of 'buoyancy' instead of considering water as a medium and the relative densities of the water and the object. Extending this line of reasoning, Peng and Knowles (1993) found that the adults were likely to describe the phenomenon of magnetism in a dispositional manner. It is worth further exploring the explanations in the physical version of the study. There was not a clear internal/external dichotomy for magnets. Interaction explanations such as "she has metal feet and there's a magnet in the skateboard" could be construed as dispositional properties for both objects. Future studies could look at other phenomena in Newtonian mechanics as discussed in Peng and Knowles where there was no difference in adults from the United States or China. The materials could also be varied in their degree of humanness. Perhaps the person/situation dichotomy does not capture how people in a correspondence-bias prone culture view magnets, even though it seems analogous at first. If two identical looking square magnets were used, it is highly likely that dispositions would be assigned to both.

There are several lines of research underway that extend the paradigm in new ways. One focuses on explaining away anomalous behavior and events. Given that children performed similarly for both deterministic and probabilistic data, there must be some process

by which children discounted the anomalous information. We conducted a study where the procedure is otherwise identical to the control conditions of Chapters 2 and 3. If Josie plays 3 out of 4 times on the scooter, the question is posed about the minority event instead: "did you see that one time, Josie didn't play on the scooter? Why do you think she did that?". Then for the other children, Josie was treated as a toy as in Chapter 3. Preliminary results suggest that children treat a single versus multiple anomalies in a different manner, and when the objects are treated as people, they tend to explain away the anomalous behavior with transient states of the person, such as "she was bored" or "she didn't feel like it that time."

Another study compares the physical magnet conditions across cultures. Research underway with our collaborators in Beijing extends the procedure described in Chapter 3 to children in China. This will provide a richer picture of age differences by domain differences by cultural differences in mapping out the relationship between causal inference and attribution.

Finally, a microgenetic study in the early stages will address the question of theory change more directly. Children at the local preschool would be divided into several groups. Two groups would hear different stories every day for two weeks about two characters in similar situations to the original study. The first group would hear stories that mimicked the Person condition of the experiments. The second group would hear stories with the same statistical properties as the Situation condition. Both groups would be prompted in ach training session to explain the characters' behavior, and then predict future behavior. Unlike the original study however, the children would then get feedback on whether their prediction was "correct" and have the opportunity to view what the two characters actually did. The dolls would always behave in a consistent matter with the person or situation condition. Then a control group of children would play unrelated games with the experimenter to control for the element of interaction. At the end of the two weeks, they would receive the "classic" version of the experiment, as in Rholes and Ruble (1984), where they are asked to generalize about a single instance of behavior. We would predict that children in the Person training group would be more likely to jump to the conclusion that he did not share his lunch because he is a mean person than children in the Situation training and control groups. And perhaps most intriguingly, we could bias children in the opposite direction in the Situation training group to overweight the context. However, it seems unlikely that even a two-week intervention could counteract the influence of the dominant culture.

This study deliberately used a fairly neutral trait pair: brave and scared. Most studies with children look at heavily valenced traits that usually fall into broader nice/mean categories. Perhaps there are different prior beliefs for different traits. Positively valenced traits can be explained away by socially sanctioned behavior. The person could simply be generous because it is socially efficacious to do so. Young children place more weight on evidence that a person is mean versus a person is nice. Children might have a stronger internal bias for antisocial behavior, as for the Rholes and Ruble (1984) scenario with the child who refuses to share. We should also adapt the study procedure to mimic other types of biases, such as the aforementioned in-group out-group bias. In this case, children in China might discount the local evidence in favor of prior beliefs about (usually out-group) group members at a much higher rate.

By applying the principles of causal inference from covariation data to the social domain, we can see how children interpret new information through the lens of their prior beliefs. We can look at whether children spontaneously explain people's behavior as something about the person or situation. Although we still do not know the exact mechanism by which children form culturally-based attributional schemas, or the extent of the ramifications of relying on those schemas, causal Bayes nets provide a useful tool for modeling the accompanying theory shift that takes place in childhood and the complex relationship between prior beliefs and evidence as we solidify our theories about people. As we discover new ways to quantify culture, we will better understand its influence on the formation of attributional theories.

References

Alvarez, J. M., Ruble, D. N., & Bolger, N. (2001). Trait understanding or evaluative reasoning? An analysis of children's behavioral predictions. *Child Development*, *72*(5), 1409–1425. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/1467-8624.00356/abstract

Baillargeon, R. (2008). Innate Ideas Revisited: For a Principle of Persistence in Infants' Physical Reasoning. *Perspectives on Psychological Science*, *3*(1), 2–13. doi:10.1111/j.1745-6916.2008.00056.x

Bartsch, K., & Wellman, H. M. (1995). Children talk about the mind. New York: Oxford University Press.

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, *68*(3), 530–543.

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, *42*(3), 500–513.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*(2), 307–324.

Buchanan, D. W., & Sobel, D. M. (2011). Mechanism-based causal reasoning in young children. *Child Development*, *82*(6), 2053–2066. doi:10.1111/j.1467-8624.2011.01646.x

Cheng, P., & Novick, L. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, *58*(4), 545–567.

Choi, I., Nisbett, R. E., & Norenzayan, A. (1999). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*, 47–63.

Dweck, C. S., Hong, Y. Y., & Chiu, C. Y. (1993). Implicit Theories Individual Differences in the Likelihood and Meaning of Dispositional Inference. *Personality and Social Psychology Bulletin*, *19*(5), 644–656. doi:10.1177/0146167293195015

Ferguson, T. J., Olthof, T., Luiten, A., & Rule, B. (1984). Children's use of observed behavioral frequency versus behavioral covariation in ascribing dispositions to others. *Child Development*, *55*(6), 2094–2105.

Flavell, J. H., Flavell, E. R., Green, F. L., & Moses, L. J. (1990). Young children's understanding of fact beliefs versus value beliefs. *Child Development*, *61*(4), 915–928.

Gelman, R. (2000). Domain specificity and variability in cognitive development. *Child Development*, *71*(4), 854–856. doi:10.2307/1132328

Gelman, S. A., & Noles, N. S. (2010). Domains and naïve theories. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 490–502. doi:10.1002/wcs.124

Gibson, E. J. (1995). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Advances in infancy research*, *9*, XXI–LXI.

Gonzalez, C. M., Zosuls, K. M., & Ruble, D. N. (2010). Traits as dimensions or categories? Developmental change in the understanding of trait terms. *Developmental Psychology; Developmental Psychology*, *46*(5), 1078–1088. doi:10.1037/a0020207

Goodman, N. (1983). Fact, fiction, and forecast. Cambridge: Harvard University Press.

Gopnik, A., & Schulz, L. (2007). Causal Learning: Psychology, Philosophy, and Computation: Psychology, Philosophy, and Computation.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review*, *111*(1), 3–32. doi:10.1037/0033-295X.111.1.3

Gopnik, A., Sobel, D. M., Schulz, L. E., & 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*(5), 620–629.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, *51*(4), 334–384.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, *116*(4), 661–716. doi:10.1037/a0017201

Hewstone, M., & Jaspars, J. (1987). Covariation and causal attribution: A logical model of the intuitive analysis of variance. *Journal of Personality and Social Psychology*, *53*(4), 663–672.

Heyman, G. D. (2009). *Children's reasoning about traits. Rational Constructivism in Cognitive Development* (Vol. 37, pp. 105–143). Elsevier. doi:10.1016/S0065-2407(09)03703-3

Heyman, G. D., & Gelman, S. A. (1999). The use of trait labels in making psychological inferences. *Child Development*, *70*(3), 604–619.

Heyman, G. D., & Gelman, S. A. (2000a). Preschool children's use of trait labels to make inductive inferences. *Journal of Experimental Child Psychology*, *77*(1), 1–19. doi:10.1006/jecp.1999.2555

Heyman, G. D., & Gelman, S. A. (2000b). Preschool children's use of novel predicates to make inductive inferences about people. *Cognitive Development*, *15*(3), 263–280.

Higgins, E. T., & Bryant, S. L. (1982). Consensus information and fundamental attribution error: The role of development and in-group versus out-group knowledge. *Journal of Personality and Social Psychology*, *43*(5), 889–900.

Hong, Y.-Y., Chiu, C. Y., Dweck, C. S., Lin, D. M., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*, 588–599.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, *3*(1), 1–24.

Jones, E. E., & Nisbett, R. E. (1971). The actor and the observer: Divergent perceptions of the causes of behavior.

Kalish, C. W. (2002). Children's predictions of consistency in people's actions. *Cognition*, *84*(3), 237–265. doi:10.1016/S0010-0277(02)00052-5

Karmiloff-Smith, A. (1994). Beyond modularity: A developmental perspective on cognitive science. *International Journal of Language & Communication Disorders*, *29*(1), 95–105.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (ed.), Nebraska Symposium on Motivation (Volume 15, pp. 192-238). Lincoln: University of Nebraska Press. doi:1968-13540-001

Kemp, C., Goodman, N. D., & Tenenbaum, J. B. (2008). Theory acquisition and the language of thought. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science, 10(3)*, 307–321. doi:10.1111/j.1467-7687.2007.00585.x

Krull, D. S., Loy, M. H. M., Lin, J., Wang, C. F., Chen, S., & Zhao, X. (1999). The Fundamental Fundamental Attribution Error: Correspondence Bias in Individualist and Collectivist Cultures. *Personality and Social Psychology Bulletin*, *25*(10), 1208–1219. doi:10.1177/0146167299258003

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*(1), 186–196. doi:10.1037/0012-1649.43.1.186

Levy, S. R., & Dweck, C. S. (1998). Trait-versus process-focused social judgment. *Social Cognition*, *16*(1), 151–172.

Levy, S. R., & Dweck, C. S. (1999). The impact of children's static versus dynamic conceptions of people on stereotype formation. *Child Development*, *70*(5), 1163–1180.

Lieberman, M. D. (2005). Attributional Inference Across Cultures: Similar Automatic Attributions and Different Controlled Corrections. *Personality and Social Psychology Bulletin*, *31*(7), 889–901. doi:10.1177/0146167204274094

Liu, D., Gelman, S. A., & Wellman, H. M. (2007). Components of young children's trait understanding: behavior-to-trait inferences and trait-to-behavior predictions. *Child Development*, *78*(5), 1543–1558. doi:10.1111/j.1467-8624.2007.01082.x

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*(2), 523–531. doi:10.1037/0012-1649.44.2.523

Lockhart, K., Nakashima, N., Inagaki, K., & Keil, F. C. (2008). From ugly duckling to swan? Japanese and American beliefs about the stability and origins of traits. *Cognitive Development*, *23*(1), 155–179.

Lucas, C. G., Gopnik, A, & Griffiths, T. L. (2010). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. In Proceedings of the 32nd Annual Conference of the Cognitive Science Society, S. Ohlsson, R. Catrambone, Eds.(Cognitive Science Society, Austin, Tex.), 28-52.

Malle, B. F. (2011). *Chapter 6 - Time to Give Up the Dogmas of Attribution:An Alternative Theory of Behavior Explanation. Advances in Experimental Social Psychology* (1st ed., Vol. 44, pp. 297–352). Elsevier Inc. doi:10.1016/B978-0-12-385522-0.00006-8

Malle, B. F., Knobe, J. M., & Nelson, S. E. (2007). Actor-observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology*, *93*(4), 491.

Matsunaga, A. (2002). Preschool children's inferences about personality traits. *Japanese Journal of Developmental Psychology*, *13*(2), 168–177.

Miller, J. G. (1984). Culture and development of everyday social explanation. *Journal of Personality and Social Psychology*, *46*(5), 961–978. doi:10.1037/0022-3514.46.5.961

Molden, D. C., Plaks, J. E., & Dweck, C. S. (2006). "Meaningful" social inferences: Effects of implicit theories on inferential processes. *Journal of Experimental Social Psychology*, *42*(6), 738–752. doi:10.1016/j.jesp.2005.11.005

Morris, M. W., & Larrick, R. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review*, *102*(2), 331–355.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*, 949–971. doi:10.1037/0022-3514.67.6.949

Na, J., & Kitayama, S. (2011). Spontaneous Trait Inference Is Culture-Specific: Behavioral and Neural Evidence. *Psychological Science*, *22*(8), 1025–1032. doi:10.1177/0956797611414727

Nisbett, R. E. (2003). The geography of thought: How Asians and Westerners think differently... and why.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*(2), 291–310. doi:10.1037//0033-295X.108.2.291

Olson, K. R., Banaji, M. R., Dweck, C. S., & Spelke, E. S. (2006). Children's Biased Evaluations of Lucky Versus Unlucky People and Their Social Groups. *Psychological Science*, *17*(10), 845–846. doi:10.1111/j.1467-9280.2006.01792.x

Orvis, B., Cunningham, J., & Kelley, H. (1975). A closer examination of causal inference: The roles of consensus, distinctiveness, and consistency information. *Journal of Personality and Social Psychology*, *32*(4), 605–616.

Patrick, B. C., Skinner, E. A., & Connell, J. P. (1993). What motivates children's behavior and emotion? Joint effects of perceived control and autonomy in the academic domain. *Journal of Personality and Social Psychology*, *65*, 781–781.

Pearl, J. (2000). Causality: models, reasoning and inference, *29*.

Peevers, B. H., & Secord, P. F. (1973). Developmental changes in attribution of descriptive concepts to persons. *Journal of Personality and Social Psychology*, *27*(1), 120–128. doi:doi:10.1037/h0034469

Peng, K., & Knowles, E. D. (2003). Culture, education, and the attribution of physical causality. *Personality and Social Psychology Bulletin*, *29*(10), 1272–1284. doi:10.1177/0146167203254601

Pettigrew, T. F. (1979). The Ultimate Attribution Error: Extending Allport's Cognitive Analysis of Prejudice. *Personality and Social Psychology Bulletin*, *5*(4), 461–476. doi:10.1177/014616727900500407

Plaks, J. E., Grant, H., & Dweck, C. S. (2005). Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception. *Journal of Personality and Social Psychology*, *88*, 245–262.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, *130*(3), 323–360.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Developmental Psychology*, *33*(1), 12.

Rholes, W., & Ruble, D. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, *55*(2), 550–560. doi:10.2307/1129966

Rosati, A. D., Knowles, E. D., Kalish, C. W., Gopnik, A., Ames, D. R., & Morris, M. W. (2001). The rocky road from acts to dispositions: Insights for attribution theory from developmental research on theories of mind. *Intentions and intentionality: Foundations of social cognition*, 287–303.

Ross, L., Amabile, T., & Steinmetz, J. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology*, *35*(7), 485–494.

Ruble, D. N., Feldman, N. S., Higgins, E. T., & Karlovac, M. (1979). Locus of causality and the use of information in the development of causal attributions1. *Journal of Personality*, *47*(4), 595–614. doi:10.1111/j.1467-6494.1979.tb00211.x

Rumelhart, D. E., & McClelland, J. L. (1987). Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Fondations, *1*.

Scholl, B. J., & Leslie, A. M. (2001). Minds, modules, and meta-analysis. *Child Development*, *72*(3), 696–701.

Schulz, L. E., & Sommerville, J. (2006). God does not play dice: Causal determinism and preschoolers' causal inferences. *Child Development*, *77*(2), 427–442.

Schulz, L. E., Bonawitz, E. B., & Griffiths, T. L. (2007). Can being scared cause tummy aches? Naive theories, ambiguous evidence, and preschoolers' causal inferences. *Developmental Psychology*, *43*(5), 1124–1139. doi:10.1037/0012-1649.43.5.1124

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, *47*(5), 1239.

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). Intergroup conflict and cooperation: The Robbers Cave experiment, *10*.

Shimizu, Y. (2000). Development of Trait Inference: Do Young Children Understand the Causal Relation of Trait, Motive, and Behavior? *Japanese Journal of Educational Psychology*, *48*(3), 255–266.

Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the society for research in child development*, *47*(1), 1–51. doi:10.2307/1165893

Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: The development of causal reasoning in toddlers and infants. *Developmental Psychology*, *42*(6), 1103–1115.

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

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

Spelke, E. (1994). Initial knowledge: six suggestions. *Cognition*, *50*(1-3), 431–445. doi:10.1016/0010-0277(94)90039-6

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605–632.

Spirtes, P., Glymour, C. N., & Scheines, R. (2000). *Causation Prediction & Search 2e*. MIT Press.

Steyvers, M. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*(3), 453–489. doi:10.1016/S0364-0213(03)00010-7

Sutton, R., & McClure, J. (2001). Covariational influences on goal-based explanation: An integrative model. *Journal of Personality and Social Psychology*, *80*(2), 222–236.

Tamir, M., John, O. P., Srivastava, S., & Gross, J. J. (2007). Implicit theories of emotion: Affective and social outcomes across a major life transition. *Journal of Personality and Social Psychology*, *92*(4), 731. doi:10.1037/0022-3514.92.4.731

Thelen, E., & Smith, L. B. (1994). A dynamic systems approach to the development of perception and action.

Waldmann, M. R., & Martignon, L. (1998). A Bayesian network model of causal learning, 1102–1107.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: foundational theories of core domains. *Annual Review of Psychology*, *43*, 337–375. doi:10.1146/annurev.ps.43.020192.002005

Wellman, H. M., & Liu, D. (2007). Causal reasoning as informed by the early development of explanations. *Causal learning: Psychology, philosophy, and computation*, 261–279. doi:10.1093/acprof:oso/9780195176803.001.0001

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science*, *17*(12), 1075–1081. doi:10.1111/j.1467-9280.2006.01830.x

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. *Psychological Science*, *18*(7), 600–606. doi:10.1111/j.1467-9280.2007.01946.x

Yong, E. (2009). East meets west: How the brain unites us all. *New Scientist, 2698*.

Appendices

Table 1: Summary of Experimental Design by Condition

	Deterministic	Doll		Probabilistic	Doll	
Study condition	Activity	A	B	Activity	A	B
Proportion of playing trials to approaches						
Person	I	4/4	0/4	I	3/4	1/4
	II	4/4	0/4	II	3/4	1/4
Situation	I	4/4	4/4	I	3/4	3/4
	II	0/4	0/4	II	1/4	1/4
Control	I	4/4	—	I	6/8	—
	II	—	0/4	II	—	2/8

Table 2: Examples of person and situation explanations

Person

Mental states
She wanted to splash
She thinks there's a shark in the water
She thinks she might fall off
She learned how to ride her bike
She was in the mood for it
She liked it
She's afraid of heights

Physical attributes
She's younger
She's bigger
She doesn't have a helmet on

Situation

Physical object
It only has two wheels
It's too fast
It looks like fun
It looked scary
It might tip over
It's not over water & it's not high
Because it's red and blue
There's netting around it

Social situation
Josie (other doll) didn't want to
Sally (other doll) played on it
Her friend did it
Her friend wasn't there