The Function of Causal Explanatory Reasoning in Children

Cristine H. Legare (chlegare@umich.edu)

Susan A. Gelman (gelman@umich.edu)

Henry M. Wellman (hmw@umich.edu)

Tamar Kushnir (tkushnir@umich.edu)

University of Michigan, Department of Psychology, 530 Church Street
Ann Arbor, MI 48109 USA

Abstract

In two studies we investigated two competing hypotheses about the function of children's explanations. Events that were inconsistent with children's prior knowledge were simultaneously contrasted with events that were consistent with children's prior knowledge. Data provide support for the thesis that inconsistent outcomes are an especially powerful trigger for children's explanations and that children provide explanations for inconsistent outcomes that refer to underlying, internal causal properties, overriding perceptual appearances.

Keywords: Causal Reasoning; Children's Explanations.

A fundamental task for all humans is explaining why things happen. Research on conceptual development indicates that even children as young as 3 years of age can use causal knowledge to make predictions (Shultz, 1982), engage in efficacious interventions (Kushnir & Gopnik, 2007; Schulz & Gopnik, 2004), and provide explanations for phenomena in the world (Wellman, Hickling, & Schult, 1997). Not only do young children frequently seek explanations by asking questions (Callanan & Oakes, 1992; Hickling & Wellman, 2001), they also construct their own explanations. This early-developing capacity for causal reasoning raises questions about the role of explanation in constructing causal systems of knowledge (Lombrozo, 2006). Although there is mounting evidence that children's causal explanations both reveal their understandings of the world and may constitute a mechanism for advancing causal learning and the acquisition of knowledge (Amsterlaw & Wellman, 2006; Bartsch & Wellman, 1989; Gopnik & Meltzoff, 1997; Siegler, 1995), direct evidence is needed concerning the role and function of children's own explanations.

Children's explanations may serve at least two distinct functions. One possibility is that children's explanations are largely confirmatory, that is, they express regularities and causal links that children have previously learned and that accord with their current knowledge. An alternative, intriguing possibility is that explanations provide children with the opportunity to generate new hypotheses regarding events that, at first, disconfirm their current knowledge. In scientific theorizing, disconfirming events play a special role in provoking explanations. Do children engage in an analogous process? Imagine that a child sees two equivalent events, one in accord with prior knowledge and the other not. If explanation is largely confirmatory, children should simply explain what they already have an explanation for. If explanation is instead responsive to discordant or anomalous information, children should explain the event that falls outside their prior knowledge or expectations. This is the scenario we use as an experimental paradigm to examine these competing hypotheses about the function of children's explanations. Moreover, we examined the nature of children's explanations, specifically whether children would provide explanations primarily in terms of surface features and past histories, or whether they would offer explanations focused on less-obvious properties.

Study 1

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

We examined which events led to (a) greater interest and attention, (b) more causal explanation-seeking, and most importantly, (c) increased explanation. If the role of explanations for children is confirmatory, children should be interested in and provide explanations for the consistent event. If "anomaly" plays a special role they could be specifically interested in and provide hypotheses and explanations for the inconsistent event.

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

Even 3-year-olds can categorize objects in terms of novel, non-obvious properties (Gelman & Coley, 1990; Gelman & Markman, 1986, 1987), apply names to objects with the same functional properties (Kemler-Nelson, 1995), and categorize and name objects based on novel causal properties (Gopnik & Sobel, 2000), overriding perceptual appearances. Whether and how preschool children might make use of this information in their causal explanations is an open question. We were interested in investigating whether children's explanations typically include information about function and causal properties. If this is the case, then it would be a strong indication that children use information about causal properties and function when reasoning about objects.

Method

Participants Sixteen 3- (*M* =3,6; range 3,2 to 3,11), 16 4- (*M*=4,6; range 4,0 to 4,11), and 16 5-year-olds (*M*=5,3; range 5,0 to 5,6), participated, including approximately equal numbers of boys and girls.

Materials Four "light boxes" (5 inches by 5 inches by 5 inches) were made of wood with thin wooden tops. The boxes were identical except that two were painted red and two yellow. Each box was connected to an electrical outlet and a switch box. Each pair of boxes (red or yellow) was attached to the same switch box and was operated surreptitiously and out-of-sight by a confederate. If the switchbox was turned on, the box would light up and stay on until it was switched off. Alternately, if the switchbox was turned off, the box would turn off and stay off until it was switched on. Each box was turned off or on as soon as an object made contact with it, and would stay off or on until the object was removed. This yielded the strong impression that some objects turned the boxes on, some turned them off, and some had no effect on the boxes. The switchbox was hidden from the children's view, and none of the participants mentioned the confederate or indicated any suspicion that the confederate influenced the functioning of the boxes. There were four perceptually distinct sets of objects, each consisting of five identical wooden blocks (20 blocks total). These were also always presented in paired sets. There was no relationship between the observed causal properties of the blocks (lighting or not lighting the boxes) and their perceptual features. Additionally, the placement of the objects on the boxes was counterbalanced in order to prevent potential effects of the perceptive shape of the objects on the probability of attracting attention and eliciting explanations.

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

In the following test trials, two contrasting objects were again simultaneously placed on the boxes. One object worked in a manner consistent with prior knowledge and the other object functioned in a manner that was inconsistent. Note that unlike the confirmation trial, in the test trials, the outcomes in both cases were perceptually identical. Both boxes lit up in one test trial and both boxes did not light up in the other test trial. Thus, to us, and perhaps to the child, some objects caused or failed to cause their normative outcome. For example, in a generative cause case (Figure 1), a starter failed to start its target device, or in an inhibitory case, a stopper did not turn off its device (Figure 2). After viewing the paired outcomes of each trial, children were asked a non-specific explanatory question, "Why did that happen?"

Figures

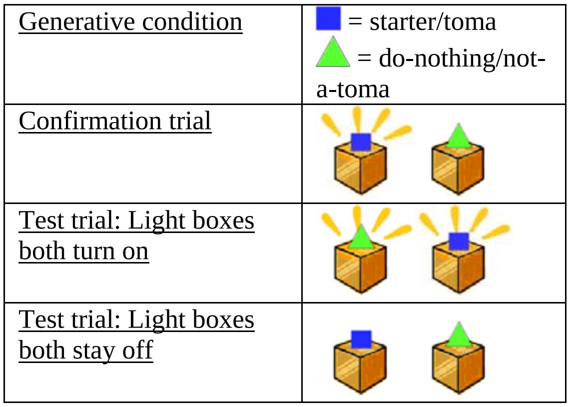

Figure 1: Studies 1 & 2: Generative condition

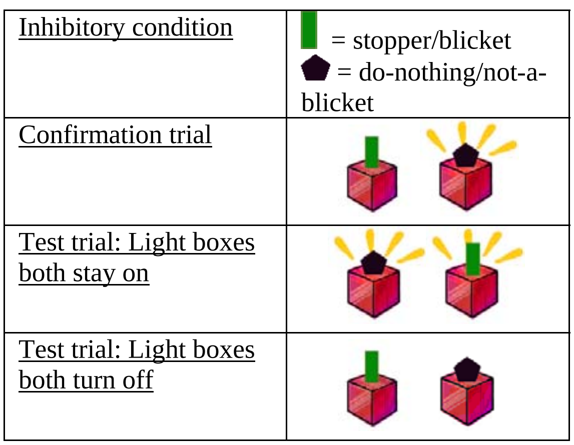

Figure 2: Studies 1 & 2: Inhibitory condition

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

During the confirmation trial, the outcome the child looked at first was coded (starter/stopper =1, do-nothing=0), followed by the outcome the child mentioned first (starter/stopper=1, do-nothing=0, both=0.5). For example, if a child referred to both outcomes in their first explanation, it was coded as 0.5. The two test trials per condition were counterbalanced; in half of the trials both of the boxes lit up first, and in half of the trials the boxes did not light up first. As in the confirmation trial, the outcome the child looked at first was coded for (inconsistent=1, consistent=0), followed by the outcome the child mentioned first (inconsistent=1, consistent=0, both=0.5). After coding whether the child provided an explanation for the inconsistent or consistent outcome (or both), the outcome the child explained first was then coded (inconsistent =1, consistent=0, both=0.5).

Explanatory Response Categories The kind of explanation provided was coded for any object that the child mentioned. Children's explanations were coded as causal or non-causal. Causal explanations for inconsistent outcomes were coded into 3 primary categories: category switch, causal function, and causal action. Children could answer the explanation question by referring to a switch in category membership based on function. For example, "It is a do-nothing now" or "It's really a starter, it only looks like a do-nothing" were coded as category switch explanations. Notice that to provide such explanations children had to ignore and go beyond perceptual identities and past history of perceptually identical objects.

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

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

Results

In the confirmation trial (out of 2 possible trials), children were more likely both to look first at the object that changed the state of the box (starter or stopper), (M=1.26), $t(47) = 2.84$, $p < .01$, and to first mention the object that changed the state of the box, (M=1.41), $t(47) = 4.57$, $p < .001$, as compared to the object that had no effect on the box (do-nothing). In contrast, in the test trials this was not the case. Instead, events that were inconsistent with children's prior expectations were very likely not only to attract children's attention but also to provoke children's explanations. Across age groups and test trials (out of 4 possible trials), children were much more likely to look first at the inconsistent outcomes than the consistent outcomes, (M=3.73), $t(47) = 18.61$, $p < .001$. Likewise, children were much more likely to explain the inconsistent outcome first, (M=3.18), $t(47) = 11.72$, $p < .001$. Additionally, overall, providing an explanation for the inconsistent outcome (M=3.98), was much more likely than providing an explanation for the consistent outcome, (M=2.56), $t(47) = 6.95$, $p < .001$.

Although there was a significant increase in causal explanations with age, $F(2,47) = 7.44$, $p < .01$, the majority of children's explanations for inconsistent outcomes at all age groups were causal (55% of the explanations provided by three-year-olds, 81% of the explanations provided by 4-year-olds, and 95% of the explanations provided by 5-year-olds).

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

Discussion

The data from Study 1 demonstrate that inconsistent outcomes are an especially powerful trigger for children's explanations. Across the generative and inhibitory conditions, children were much more likely to attend to and explain the outcome that was inconsistent with their prior knowledge. This is significant for several reasons. If children had just been paying attention to the perceptual outcome (the boxes turning off or on), the outcome they explained should have been at chance, given that in the test trials the outcomes were identical (both boxes were on or off). Additionally, given that children already had an explanation for the consistent outcome, they could have easily ignored the inconsistent outcome (for which they were given no explanation) and just explained the consistent outcome. However, this was not the case. Thus children's explanations were not merely confirmatory; inconsistent outcomes were highly noteworthy to preschool children and they provided explanations for such outcomes that referred to underlying, internal causal properties, overriding perceptual appearances.

Many of children's explanations referred to the functioning of the objects or light boxes in some fashion. This was expected, but additionally some responses might have been especially scaffolded and provoked in Study 1 because the object names (starters, stoppers, and do-nothings) were straight-forward descriptions of their functions. In addition, explanations of events not in accord with prior knowledge could have been specifically scaffolded by such terms because the objects then violated their labeled functions. Children's focus and explanations were informative nonetheless, but it is important to consider this alternative hypothesis. In order to investigate this possibility and to determine if the same effects are found without this cue, Study 2 was designed to examine the effect of novel, non-functional labels for the objects (toma, blicket) and their influence on (a) the kinds of outcomes (consistent or inconsistent) children attend to and explain and (b) the kinds of explanations children provide for these outcomes.

Study 2

Children's causal explanations may be especially revealing of the nature of children's object concepts by (a) privileging function information over perceptual features, (b) evoking causal mechanisms and internal, underlying causal properties, and (c) spontaneously utilizing novel labels to redefine category boundaries around shared function. Given that children can extend novel labels to new categories based on function, we predict that as in Study 1, children will focus more on explaining events inconsistent with prior knowledge, even when the labels themselves do not carry such functional expectations. Further, information about function and underlying causal properties will be found in children's causal explanations, even when the labels themselves are novel and do not, in themselves, provide additional causal information.

Method

Participants Sixteen 3- (M= 3,5; range 3,0 to 3,10), and 16 4-year-olds (M=4,4; range 4,0 to 4,11) participated, including approximately equal numbers of boys and girls. None of the participants in Study 2 participated in Study 1.

Materials The materials used in Study 2 were identical to Study 1.

Procedure The procedure for Study 2 was modeled on Study 1, with a critical modification: the objects were given non-functional, novel labels (“toma” / “not a toma” and a “blicket” / “not a blicket”). The confirmation and test trials were identical to Study 1.

Transcribing and Coding The same procedure for coding Study 1 was used to code Study 2.

Results

As in Study 1, inconsistent outcomes not only attracted children’s attention but also provoked children’s explanations. Across age groups and test trials (ranging from 0-4), children were much more likely to look first at the inconsistent outcomes, than the consistent outcomes, (M=3.25), $t(31) = 7.44$, $p < .001$. Likewise, children were much more likely to explain the inconsistent outcome first, (M=3.05), $t(31) = 9.86$, $p < .001$. Additionally, providing an explanation for the inconsistent outcome, (M=3.94), was more likely than providing an explanation for the consistent outcome, (M=2.88), $t(31) = 5.17$, $p < .001$.

In the confirmation trial (ranging from 0-2), although children were more likely to look first at the object that changed the state of the box (started or stopped it), (M=1.28), $t(31) = 2.06$, $p < .05$, they were equally likely to mention the object that changed the state of the box as the object that had no effect on the box, (M=1.23), $t(31) = 1.97$, *ns*.

The majority of children’s explanations for inconsistent outcomes were causal (81% of the explanations provided by three-year-olds, and 84% of the explanations provided by 4-year-olds). However, unlike in Study 1, 3-year-olds were just as likely as 4-year-olds to give causal explanations, $F(1,31) = .16$, *ns*. This indicates that children’s explanations referring to the function of the objects cannot simply be attributed to the functional labels used in Study 1.

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

Discussion

Study 2 provides further support for the hypothesis that inconsistent outcomes are an especially powerful trigger for children’s explanations. As in Study 1, across the generative and inhibitory conditions, children were much more likely to attend to and explain the inconsistent outcome, overriding perceptual appearances in favor of function information, even without the benefit of functional labels. In fact, children provided more causal function explanations in Study 2 than in Study 1. Our data demonstrate that even children as young as three-years-old use information about underlying causal properties to re-define category membership in their explanations, even when the labels are novel and not tied to function.

General Discussion

In order to address two competing hypotheses about the function of children’s own explanations, events that were consistent with children’s prior knowledge were contrasted with events that were inconsistent with prior knowledge. Children’s explanations were the primary dependent measure.

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

Furthermore, children’s explanations for inconsistent outcomes referred to underlying, internal causal properties, overriding perceptual appearances. These data suggest that children are looking for explanations that extend beyond the available evidence. Furthermore, children’s explanations may play an active, important role in the learning process. If children use explanation as a mechanism for acquiring new knowledge, they should provide explanations for events that have the potential to teach them something new. Explaining inconsistent events provides just such an opportunity, and such explanatory biases might aid in learning. Although our data do not speak to this question directly, an interesting direction for future research would be to directly examine the influence of children’s explanations on learning and knowledge acquisition.

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

Acknowledgements

This research was supported by a grant from the McDonnell Foundation, by NICHD grant HD22149 to Wellman, and NICHD grant HD36043 to Gelman. We are grateful to the teachers, staff, and children of the University of Michigan Children's Centers, Pooh Corner Preschool, and Bemis Farms Preschool for participating in this research. We would also like to thank Andrea Kiss for her able assistance with data collection. Address for correspondence: Cristine H. Legare, Department of Psychology, 530 Church St., University of Michigan, Ann Arbor, MI 48109-1043. E-mail: chlegare@umich.edu.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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