

RUNNING HEAD: Pedagogy and counterevidence

**Look again: Pedagogical demonstration facilitates children's use of counterevidence**

Lucas P. Butler

University of Maryland

Submitted to *Child Development* on November 1, 2017

Author Note

Lucas P. Butler, Department of Human Development and Quantitative Methodology,  
University of Maryland, College Park.

We thank Hailey Gibbs, Ariel Horn, and Erica Zippert for assistance with data collection and coding, and Andrei Cimpian for helpful feedback on a previous version of this manuscript.

Correspondence concerning this article should be addressed to Lucas P. Butler,  
Department of Human Development and Quantitative Methodology, University of Maryland,  
3942 Campus Drive, College Park, MD 20742. Contact: [lpbutler@umd.edu](mailto:lpbutler@umd.edu).

### **Abstract**

In learning about the world, we must not only make inferences based on minimal evidence, but must deal with conflicting evidence and question those initial inferences when they appear to be wrong. In three experiments (N=96), we found that in some cases young children only revise their causal beliefs when conflicting evidence is explicitly demonstrated for them. Four- and 5-year-old children inferred a rule about what objects had causal powers, and then saw evidence conflicting with that initial inference. Critically, the conflicting evidence was produced either instrumentally and intentionally, or demonstrated communicatively and pedagogically. Only when evidence was explicitly demonstrated for them did children revise their initial hypothesis and use a subtle clue to infer the correct rule.

*Keywords:* belief revision; categorization; pedagogy; social cognition; conceptual development

### **Look again: Pedagogical demonstration facilitates children's use of counterevidence**

As humans, we live in a world of unfathomable uncertainty, yet assume that within that noise there are signals that will help us derive meaning and order. We search for meaningful patterns that help explain the structure of the world (Cimpian & Salomon, 2014), and we both strive to and succeed in identifying those patterns and using them to organize our knowledge, actions, judgments, and decisions across a wide variety of contexts (Gelman, 2003; Gelman & Wellman, 1991; Keil, 1989; Markman, 1989). And yet, as should be expected when drawing uncertain inferences from minimal data, much of the time the initial inferences that we make are wrong. We infer that the earth is flat, that the sun orbits the earth, that invisible entities determine everyday events, that moving objects have a hidden force inside them.

Given that the evidence available to us is insufficient to ensure accuracy, and frequently leads to incorrect conclusions and common misconceptions, the capacity to revise our initial inferences based on of new and conflicting evidence is critical to constructing a coherent understanding of the world. And though this is undoubtedly true throughout the lifespan, it is especially true for children forming the initial causal theories that will frame their subsequent learning. Even very young children are sensitive to surprising or unexpected evidence, at least when it presents a clear conflict (Legare, 2012; Stahl & Feigenson, 2015). But relevant counterevidence may often be subtle and easily overlooked unless there is reason to think it might be particularly important. In making inferences, children use others as a source of knowledge to help guide inductive inferences in the face of uncertainty and to gauge the importance of new information. The current research aims to bring this work to bear on the critical process of hypothesis revision in the face of counterevidence. We asked whether children

might be more likely to make use of counterevidence in revising their initial hypotheses, when that counterevidence was deliberately demonstrated for their pedagogical benefit, compared with seeing the identical counterevidence produced in an intentional, but instrumental and non-pedagogical manner.

In the following sections, I first review the literature on the role that social cues play in children's inductive inference process. Then, I turn to the literature on children's sensitivity to conflicting or surprising evidence. Finally, I motivate the current studies, which aim to integrate these two literatures in order to address how social cues might facilitate children's reasoning from counterevidence.

### **Children's use of social cues to guide inductive inferences**

A foundational capacity for learning is the capacity to make broad inductive generalizations on the basis of limited evidence. This capacity is central to causal reasoning, is necessary for prediction, explanation, and underlies conceptual development more broadly. But inference from minimal evidence presents an inductive problem, in which children need to figure out what of all the novel information they encounter provides good bases for generalization, and what information should be dismissed or ignored as irrelevant, as well as the scope of the generalization that should be made if one is warranted (Goodman, 1965). Although our cognitive architecture appears to rely heavily on reasoning in terms of kinds and categories (Cimpian, 2016; Gelman, 2003; Markman, 1989), such reasoning may be insufficient to gauge, in the moment and on its own, the importance of a novel piece of minimal evidence.

Recent research has shed light on children's ability to use others as a source of knowledge to help guide such inductive inferences in the face of uncertainty. Specifically, this research has explored the ways in which children employ their social cognitive skills—their

ability to understand the intentions behind others' actions—in order to gauge the importance of information being conveyed by knowledgeable others. By early childhood, children make use of a variety of social cues, both linguistic and non-linguistic, to guide this process.

One such social cue is the kind of language others employ in conveying information. Specifically, they are highly sensitive to generic language (e.g., “*birds* have rocks in their stomachs,” as opposed to “*this bird* has rocks in his stomach”) that links knowledge directly to the kind. Generic language leads children to view properties as relevant, important, and generalizable to the categories to which utterances refer (Prasada, 2000; Prasada & Dillingham, 2006; Gelman, Star, and Flukes 2002; Cimpian & Cadena, 2010; Cimpian & Erickson, 2012; Cimpian & Markman, 2009, 2011; Cimpian, Mu, & Erickson, 2012; Hollander, Gelman, & Raman, 2009). And specifically framing a causal or statistical problem appears to make children more sensitive to the evidence relevant to answering that problem (Butler & Markman, 2012a; Garvin & Woodward, 2015).

Children are also highly sensitive to non-linguistic cues that may indicate the importance or generalizability of new information in a way analogous to generic language, in particular social cues such as eye gaze and joint attention that signal a pedagogical intention to communicate relevant information (Csibra & Shamsuddeen, 2015). Seeing novel properties or actions demonstrated intentionally, with cues suggesting the pedagogical intent to transmit important information, leads children to make inductive generalizations about that evidence that are stronger and more resistant to counter-evidence (Butler & Markman, 2012b, 2016; Butler & Tomasello, 2016; Hernik & Csibra, 2015), leads them to construct radically different conceptions of novel categories (Butler & Markman, 2014; Futó, Téglás, Csibra, & Gergely, 2010; Kovács, Téglás, Gergely, & Csibra, 2016; Yoon, Johnson, & Csibra, 2008), and leads them to infer that

the demonstrated action is the only (Bonawitz, Shafto et al., 2011) or normatively correct (Schmidt, Butler, Heinz, & Tomasello, 2016; Vrendenburgh, Kushnir, & Casasola, 2014) way to act. Taken together, this evidence suggests that children's sensitivity to whether or not actions are carried out with pedagogical intent towards the child shapes the inferences children make about the evidence those actions produce.

### **Sensitivity to counterevidence in inductive inference**

The research described above paints a clear picture of the role that social cues play in children's inductive inference process. However, all of this research focuses on children's initial inductive inferences—their construction of knowledge from nothing to an initial hypothesis or implicit theory. But as we describe above our initial theories about the world are often wrong. Just as important as the ability to make initial inferences on the basis of minimal evidence is the capacity to evaluate and revise those inferences in light of new evidence. This was of course one of the lynchpins of Piaget's (1952) theory of cognitive development, in which new information leads to disequilibrium, which then must be resolved either through assimilation of that evidence into existing schemas, or through accommodation of the schema to fit the new evidence.

We know from decades of research with adults that humans are famously bad at reasoning about evidence that conflicts with their beliefs (see Nickerson, 1998 for a thorough review). Philosophers as early as Francis Bacon (1620/1939) have noted that the human mind, once decided on an opinion or conception of something, tends to search for information confirmation that opinion or conception and distorts or ignores information that conflicts with it. Classic work in cognitive psychology (e.g., Bruner, Goodnow, & Austin, 1956; Wason, 1960) has shown repeatedly that people often test hypotheses by choosing only examples that would fit

their hypothesis. And such patterns are not merely limited to physical reasoning, but to social reasoning as well (e.g., Nisbett & Ross, 1980).

What about children? At least when evidence presents a stark conflict with what their prior knowledge and understanding of how the world works, children show a motivation to engage with and explore counterevidence to try to resolve that conflict. For example, infants shown events that conflict with their intuitive expectations about how objects should behave in the physical world (e.g., a car that hovers in mid-air or seems to magically pass through a solid wall), not only show surprise at such events but selectively explore those objects in order to make sure the objects truly behave in such a counter-intuitive way. Likewise, preschoolers shown evidence that an object does not balance the way they would expect it to selectively explore to figure out why (Bonawitz, Ullman, Gopnik, & Tenenbaum, 2012). Finally, preschoolers who are explicitly taught a causal rule (e.g., that “blickets” make a machine turn on, while “daxes” do not), and then shown evidence that conflicts with that rule (e.g., a “dax” that works and a “blicket” that does not) selectively try to explain and explore that inconsistency in the hopes of understanding it (Legare, 2012; Legare et al., 2010). This is not to suggest that young children are immune from confirmation bias—certainly that seems unlikely at best—but rather to show that in cases where a conflict is clear and stark for one reason or another, children are motivated to and can successfully explore to resolve the conflict.

### **The current research**

The current research aims to bring together the literatures on social cues in inductive inference and children’s reasoning about counterevidence. Specifically, we note that previous research has largely overlooked the issue of what leads children to recognize that a conflict is relevant to their ongoing process of hypothesis formation and revision, and that they ought to

search for additional evidence that might help them resolve that conflict. In prior work, the conflicts have been quite stark, either because the evidence conflicted with children's intuitive conceptions of the physical world (Bonawitz et al., 2012; Stahl & Feigenson, 2015, 2017), or because the initial rule was explicitly taught and conflict was made vivid through deliberate demonstration or prompts for explanation (Legare, 2012; Legare et al., 2010).

But in many cases, evidence may not present such a stark conflict. Clearly evidence that conflicts with the laws of the natural world is rare. Evidence that conflicts with children's intuitive conceptions of the physical world may be more common, but it may also be easy to overlook or dismiss until enough of it has accumulated to force belief revision. And adults are not always teaching causal rules or clearly exhibiting conflicting cases. There are thus many cases in which conflicts may easily be overlooked, either because the conflict is not so stark or because the evidence is easily dismissed for various reasons.

This seems like a natural place where social cues might help close the gap. In the current research, we ask whether children might be more likely to make use of counterevidence in exploring and revising their initial hypotheses, when that counterevidence has been pedagogically demonstrated for their benefit. Specifically, we manipulate whether or not counterevidence is demonstrated pedagogically or produced, in an intentional, instrumental, but non-pedagogical manner. Children then see a potential clue that might resolve the conflict, and we ask whether children will use that clue accordingly to revise their initial hypothesis.

## **Experiment 1**

### **Method**



The research described below and in Experiments 2 and 3 was approved by our university's Institutional Review Board. Informed consent was obtained from parents or legal guardians for children's participation, and verbal assent was given by participating children.

**Participants.** Thirty-two preschoolers (Mean age = 53 months; 16 girls) participated. Children were recruited at a children's museum in a mid-sized city in the mid-Atlantic region. Children were predominately middle- and upper-middle-class, but represented diverse ethnic and racial backgrounds. Four additional children were testing but excluded from the final sample, either because of experiment error ( $n=1$ ) or because they did not want to explore in the initial exploration phase ( $n=3$ ).

**Materials.** The materials used consisted of a novel music box toy and 8 exploration objects, 8 categorization objects, and 1 demonstration object. The music box was made from a small plastic bin (13" x 9" x 6") covered in red duct tape and blue electrical tape. The top of the box was covered with black foam core and had a piece of curved white PVC piping extending from it (see Figure 1 for a picture of the experimental materials). Inside the music box was a doorbell receiver which could be activated using a wireless remote. Of the 8 exploration objects, 4 were cylinders made from PVC piping and covered in red electrical tape, and 4 were rectangles made from a wooden block and covered in red electrical tape. All of the cylinders in the exploration set had a small barcode on the bottom, made using white electrical tape and black stripes drawn with black permanent marker. None of the rectangles had this barcode. The set categorization objects was identical to the exploration objects, except that now half of each shape had the barcode on the bottom, and half did not. This was done so that barcode and shape were confounded in the exploration phase, in case children discovered the barcode during exploration

(only 2 children did, 1 in each condition). Finally, the demonstration object was an additional rectangular object with a barcode on the bottom.

[INSERT FIGURE 1 HERE]

**Procedure.** Children were invited into a private area on the museum floor. They were first introduced to a novel toy and told that the toy played music when you put certain things on it. They were then given the toy and a bin with the 8 exploration objects, and were told they could “go ahead and play.” The experimenter turned to the side, not looking at the child, and allowed the child to explore freely for 60 s. She surreptitiously kept an eye on the child’s exploration, pressing the remote for the hidden doorbell inside the toy whenever the child put one of the *active* objects (counterbalanced for which shape was active and which was inert) on the toy. This was done to give children the opportunity to discover that all objects of one shape made the toy play music, while none of a different shape did.

After children explored, the experimenter put on the demonstration object—a new rectangular identical to those that previously had failed to activate the toy—but this time the toy played music. Critically, the experimenter produced this counterevidence in two perceptually identical, but socially distinct, manners. In the pedagogical condition, the experimenter established eye contact with the child, saying, “Look, watch,” before placing the object on the toy. In the intentional condition, the experimenter placed the object on the toy without looking at or talking to the child. In both conditions, the experimenter then placed the object on its side, with the bottom facing towards the child, thus making a previously hidden barcode visible to the child. This presented a potential explanation, namely that the barcode, and not the shape,

determined each object's causal power. Importantly, the experimenter never pointed out or remarked on the barcode.

Finally, the experimenter took the music box toy away, and presented the set of 8 categorization objects to the child. As mentioned above, only half of each shape had a barcode in the bottom, but children could only discover this by looking. The experimenter asked the child to put all the objects they thought would make the toy play music in one bin, and all the objects they thought would not play music in the other bin, and looked away while children sorted the objects.

## **Results**

We first coded to make sure that children were attending to the action producing counterevidence in both conditions. All 32 children were looking either at the object or at the experimenter during the key moments during which the experimenter picked up the demonstration object, placed it on the machine, and then placed it to the side.

Our key dependent measure was whether children based their categorization judgments on object shape or presence of the hidden barcode. If children turned over at least 6 out of the 8 categorization objects to check for the barcode, and sorted them into the two bins on that basis, they were coded as categorizing according to barcode. If children sorted at least 6 out of 8 of the categorization objects according to whether or not they were the same shape as the demonstration object, they were coded as categorizing according to shape. All 32 children categorized the objects according to either shape or barcode on this criterion.

In the intentional condition, children never used the barcode to make their judgments. In contrast, in the pedagogical condition 7 out of 16 children based their judgments on presence of

the barcode, suggesting they had revised their previous causal inferences (Fisher's exact test,  $p = .007$ ; Figure 2).

[INSERT FIGURE 2 HERE]

Thus simply pedagogically demonstrating a subtle piece of counterevidence led nearly half of the children to notice and make use of a potential clue that might provide a resolution, and to use the evidence that that clue led them to generate in order to revise their initial causal hypothesis. This is a particularly striking effect given that their initial causal hypothesis was formed via active, hands-on exploration, and children have previously been shown to be especially prone to privilege evidence they have generated themselves in making causal inferences (Kushnir, Wellman, & Gelman, 2009).

## **Experiment 2**

The results of Experiment 1 suggest that even a simple pedagogical demonstration, which may draw attention to the surprising counterevidence, but not to the potential clue to resolving the conflict, has a clear effect on children's hypothesis revision. Absent such a pedagogical demonstration, when children see a clear intentional action producing conflicting evidence, but that evidence is not demonstrated explicitly for their benefit, they seem to ignore it and fail to revise their initial hypothesis. In contrast, simply a brief pedagogical clue—establishing eye contact with the child and suggesting they watch what is going to happen next—is sufficient for nearly half of the children to both pick up on the conflict and, most critically, then use the potential clue to revise their initial hypothesis.

In Experiment 2, we aimed first to replicate this initial finding with a new sample of children, as well as to push this finding one step further by asking whether children might show the same effect when the counterevidence was a *failure* for an object to have an expected causal power, rather than a surprising *success* for an object to have an unpredicted one. We hypothesized that this might be harder for children, both because the event producing the relevant counterevidence was essentially a non-event—no causal effect occurs—as well as because the potential clue—the barcode on the bottom of the demonstration object—was potentially an inhibitory cause rather than a generative one. Nevertheless, we predicted that children would show the same general effect, being more likely to revise their initial hypothesis only when they saw the relevant counterevidence demonstrated pedagogically for them.

### **Method**

**Participants.** An additional 32 preschoolers (Mean age = 53 months; 16 girls) participated. Children were recruited at the same children’s museum as in Experiment 1. As in Experiment 1, children were predominately middle- and upper-middle-class, but represented diverse ethnic and racial backgrounds. Five additional children were tested but excluded from the final sample because of experimenter error ( $n=3$ ) or because they did not want to explore in the initial exploration phase ( $n=2$ ).

**Materials.** The materials were identical to those used in Experiment 1, except that in Experiment 2 the rectangular objects in the exploration phase turned the machine on, while the cylinders did not.

**Procedure.** The procedure in Experiment 2 was identical to that used in Experiment 1, with one key change. Instead of the counterevidence being positive (i.e., an object that children would not have expected to turn the machine on did), the counterevidence was negative. That is,

the demonstration object was identical to those that children had previously discovered made the machine turn on, but in this case it failed to do so. All other aspects of the procedure remained the same.

## **Results**

As in Experiment 1, the key dependent measure was whether children based their categorization judgments on object shape or presence of the hidden barcode. In the intentional condition, children never used the barcode to make their judgments. In contrast, in the pedagogical condition 6 out of 16 children based their judgments on presence of the barcode, suggesting they had revised their previous causal inferences (Fisher's exact test,  $p = .018$ ). This replicated the finding from Experiment 1 that simply having the counterevidence demonstrated pedagogically led children to be significantly more likely to use that counterevidence to revise their initial causal hypothesis.

## **Discussion**

As in Experiment 1, children only used counterevidence to revise their initial inferences when that counterevidence was demonstrated pedagogically for them. There are, of course, several possible mechanisms may underlie these findings. One is that children simply failed to notice the counterevidence in the instrumental condition, either because they're not paying attention or because they are not attending in the right way. This seems unlikely given that all the children were looking during the demonstration. Further, children clearly recognized the causal status of the object used in the demonstration phase to produce the conflicting evidence. For example, in Experiment 1, 13 of 16 children in the intentional condition, who all categorized according to shape, nevertheless included the demonstration object in the "active" bin,

suggesting they noticed that, unlike the others of that shape, it had the same causal power as the other shape. Nevertheless, they failed to revise their causal inference.

The second alternative is that children noticed the counterevidence in the instrumental condition, but failed to see it as important or relevant to their hypothesis formation and testing process. On this account, the pedagogical demonstration led children to see the counterevidence as particularly relevant and important, and not easy to dismiss. That is, without the pedagogical cue they noticed the counterevidence, but dismissed it once they were tasked with making a causal judgment. In Experiment 3, we tested these alternatives directly.

### **Experiment 3**

Experiment 3 was designed to determine whether children's behavior in the previous two experiments was due to lack of attention or noticing of either the counterevidence or the barcode clue, or whether they in fact noticed the evidence and clue but failed to view it as *relevant* when they made their causal predictions.

Instead of giving children an explicit categorization task, in the test phase the experimenter instead gave children the same set of objects as in Experiments 1 and 2 (4 of each shape, half of each with a barcode), and let them play freely. The two alternative mechanisms described above predict distinct behavior in this case. On the first alternative, if children in the instrumental condition simply did not notice the counterevidence, or did not notice the barcode on the bottom of the demonstration object, then they might simply do nothing or play in a non-functional manner when allowed to freely explore the objects. On the second alternative, if children were equally likely to notice and attend to the counterevidence, as well as the potential clue of the barcode on the bottom of the demonstration object, and simply failed to use the counterevidence to revise their beliefs, then we would expect that in this case they would be

equally likely to curiously explore the additional objects to see if they also had the barcode, regardless of condition. That is, we predicted that children would show equal levels of spontaneous exploration of the objects in both conditions.

## **Method**

### **Participants.**

An additional 32 preschoolers (Mean age = 57 months; 16 girls) participated. Children were recruited at the same children's museum as in Experiments 1 and 2. As previously, children were predominately middle- and upper-middle-class, but represented diverse ethnic and racial backgrounds. Ten additional children were tested but excluded from the final sample because of technical problems with the music box ( $n=6$ ) or because they wanted to stop before ending the experiment ( $n=4$ ).

### **Materials.**

The materials were identical to those used in Experiment 1, except with no plastic bins for the categorization task. Additionally, technical issues with the music box led us to change the mechanism to an iPhone-activated song instead of a doorbell.

### **Procedure.**

The procedure was identical to that used in Experiment 1, with one key change. Instead of asking the children to sort the categorization objects according to which ones they thought would make the toy play music, they were simply told, "I need to go write something down for a minute, but you can go ahead and play." The experimenter then gestured in an open and welcoming way to the categorization objects, indicating the children were welcome to play with them, and then turned and pretended to write on her clipboard for 60 s.

### **Results.**



As previously, we coded to ensure that all the children were attending during the key demonstration. Our key dependent measure in Experiment 3 was not how the children sorted the objects, but whether or not they checked to see if additional objects (beyond the demonstration object) had a barcode on the bottom. Checking was operationalized as deliberately turning over an object and looking at the bottom.

The results were consistent with the first alternative, confirming this prediction. Children were equally likely to spontaneously check the objects for a barcode (6 of 16 in the pedagogical condition; 7 of 16 in the instrumental condition; Fisher's exact,  $p=1.00$ ). This suggests that the findings of Experiments 1 and 2 were not due to lack of attention to or noticing of either the conflicting evidence or of the potential clue, but rather the fact that the pedagogical demonstration changed *how* they treated that evidence—specifically as something that ought to lead them to question and revisit their initial hypothesis and search for additional evidence.

## **General Discussion**

In forming accurate causal representations of the world, we need both to be able to make broad inductive inferences on the basis of minimal evidence, but also to be able to revise those inferences on the basis of new evidence. Recent research has shown that young children are sensitive to when evidence presents a stark conflict with their expectations, and selectively explore to resolve those inconsistencies (Legare, 2012; Legare, Wellman, & Gelman, 2010; Stahl & Feigenson, 2015, 2017). The current research focuses on whether, in cases where a conflict may be easily overlooked or dismissed, pedagogical demonstration of conflicting evidence may facilitate children's ability to use it to revise their beliefs.

Indeed, simply noticing inconsistent evidence does not always lead children or adults to revise their hypotheses, as has been amply demonstrated (e.g., Dunbar & Klahr, 1988; Kuhn,

1989; Vosniadou & Brewer, 1992). This may be because individuals ignore or reject inconsistent evidence (see Chinn & Brewer, 1993), or it may be because they simply do not recognize the evidence as directly *relevant* to the hypothesis or theory they are currently engaging with. This issue of relevance is where pedagogical demonstration may have the potential to play an important role. Pedagogical demonstration is an inherently communicative act. And as argued by Sperber and Wilson (1986), intentional communication necessarily conveys relevance to the recipient of a communicative act. Building on this insight, Csibra, Gergely, and their colleagues (Csibra, 2010; Csibra & Gergely, 2009; Csibra & Shamsuddeen, 2015; Gergely & Jacob, 2012) have argued that humans are evolutionarily adapted to recognize cues to intentional communication. Moreover, they argue that young children are biased to expect information that is intentionally communicated to them to be relevant, important, and generalizable. Indeed, a number of empirical studies have shown that even infants process information differently and form different expectations when evidence is deliberately manifested for them (Futó et al., 2010; Egyed, Király, & Gergely, 2013; Yoon, Johnson, & Csibra, 2008;) and by early childhood they treat that information as more important and generalizable in a variety of ways (Bonawitz et al., 2011; Butler & Markman, 2012, 2014, 2016; Butler et al., 2015; Butler & Tomasello, 2016; Schmidt et al., 2016; Vredenburgh et al., 2015).

Building on this literature, the current study applies this framework not to children's initial inferences, but to the importance process of revising those initial inferences on the basis of subtle counterevidence. Absent pedagogical demonstration, children may not have viewed a subtle piece of conflicting evidence as relevant to their working hypothesis, and failed to revise that hypothesis in light of it. In contrast, simply having a piece of conflicting evidence pedagogically demonstrated led a large proportion of children to revise their hypotheses.

As discussed above, these findings could conceivably be based on two potential mechanisms—one attentional, and one inferential. The attentional account—that children simply failed to notice either the conflict itself or the presence of a potential clue—is inconsistent with the results of Experiment 3, in which children left to freely play with the objects were just as likely in both conditions to play with, discover, and explore the barcodes on the bottom of the categorization objects. Thus being asked to make a causal prediction, as in the first two experiments, may have led children to more explicitly consider the relevance of the counterevidence with respect to inferring causal powers. When that counterevidence was explicitly demonstrated by a knowledgeable adult, children reasoned that it must be relevant and thus were more likely to search for an explanation for its relevance and discover the hidden pattern of barcodes that explained which objects had causal powers.

The current studies tackle an important problem—how subtle differences in the manner in which adults point out or demonstrate conflicting evidence may facilitate children’s ability to use counterevidence to revise their causal beliefs. This work opens up important avenues for future research into how subtle social and contextual factors may influence children’s hypothesis formation and revision, and how this may play out in the wide variety of settings in which learning and development take place.

## References

- Bacon, F. (1939). *Novum organum*. In Burt, E. A. (Ed.), *The English philosophers from Bacon to Mill* (pp. 24-123). New York: Random House. (Original work published in 1620)
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, 120, 322–330.
- Bonawitz, E., Ullman, T., Gopnik, A., & Tenenbaum, J. (2012, November). Sticking to the Evidence? A computational and behavioral case study of micro-theory change in the domain of magnetism. In *Development and Learning and Epigenetic Robotics (ICDL), 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics* (pp. 1-6). IEEE.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.
- Butler, L. P., & Markman, E. M. (2012a). Finding the cause: Verbal framing helps children extract causal evidence embedded in a complex scene. *Journal of Cognition & Development*, 13, 38-66.
- Butler, L. P., & Markman, E. M. (2012b). Preschoolers use intentional and pedagogical cues to guide inductive inferences and exploration. *Child Development*, 83, 1416–1428.
- Butler, L. P., & Markman, E. M. (2014). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition*, 130, 116–127.
- Butler, L. P., & Markman, E. M. (2016). Navigating pedagogy: Children’s developing capacities for learning from pedagogical interactions. *Cognitive Development*, 38, 27-35.

- Butler, L. P., & Tomasello, M. (2016). Two-and 3-year-olds integrate linguistic and pedagogical cues in guiding inductive generalization and exploration. *Journal of Experimental Child Psychology, 145*, 64-78.
- Butler, L. P., Schmidt, M. F., Bürgel, J., & Tomasello, M. (2015). Young children use pedagogical cues to modulate the strength of normative inferences. *British Journal of Developmental Psychology, 33*, 476-488.
- Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1-49.
- Cimpian, A. (2016). The privileged status of category representations in early development. *Child Development Perspectives, 10*, 99–104.
- Cimpian, A., & Cadena, C. (2010). Why are dunkels sticky? Preschoolers infer functionality and intentional creation for artifact properties learned from generic language. *Cognition, 117*, 62– 68.
- Cimpian, A., & Erickson, L. C. (2012). The effect of generic statements on children's causal attributions: Questions of mechanism. *Developmental Psychology, 48*, 159–170.
- Cimpian, A., & Markman, E. M. (2009). Information learned from generic language becomes central to children's biological concepts: Evidence from their open-ended explanations. *Cognition, 113*, 14–25.
- Cimpian, A., & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child Development, 82*, 471–492.

- Cimpian, A., Mu, Y., & Erickson, L. C. (2012). Who is good at this game? Linking an activity to a social category undermines children's achievement. *Psychological Science*, 23, 533–541.
- Cimpian, A., & Salomon, E. (2014). The inference heuristic: An intuitive means of making sense of the world, and a potential precursor to psychological essentialism. *Behavioral and Brain Sciences*, 37, 461–480.
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25, 141–168.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148–153.
- Csibra, G., & Shamsudheen, R. (2015). Nonverbal generics: Human infants interpret objects as symbols of object kinds. *Annual Review of Psychology*, 66, 689–710.  
<http://doi.org/10.1146/annurev-psych-010814-015232>
- Dunbar, K., & Klahr, D. (1988). Developmental differences in scientific discovery strategies. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon. Proceedings of the 21st Carnegie-Mellon Symposium on Cognition* (pp. 109–144). Hillsdale, NJ: Erlbaum.
- Egyed, K., Király, I., & Gergely, G. (2013). Communicating shared knowledge in infancy. *Psychological Science*, 24, 1348–1353.
- Futó, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative Function Demonstration induces kind-based artifact representation in preverbal infants. *Cognition*, 117, 1–8.
- Garvin, L. E., & Woodward, A. L. (2015). Verbal framing of statistical evidence drives children's preference inferences. *Cognition*, 138, 35–48.

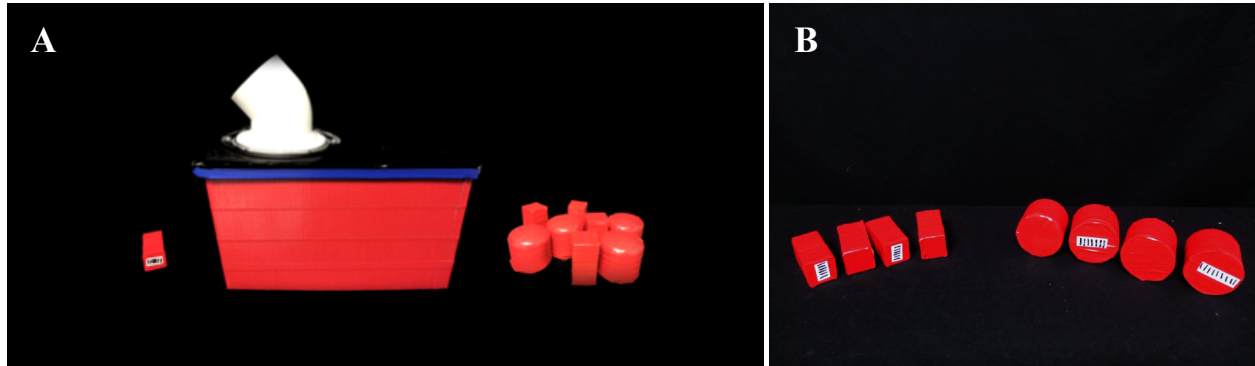
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.
- Gelman, S. A., Star, J. R., & Flukes, J. (2002). Children's use of generics in inductive inferences. *Journal of Cognition and Development*, 3, 179-199.
- Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, 38(3), 213-244.
- Gergely, G., & Jacob, P. (2012). Reasoning about instrumental and communicative agency in human infancy. In J. B. Benson (Serial Ed.) & F. Xu & T. Kushnir (Vol. Eds.), *Rational Constructivism in Cognitive Development* (pp. 59-94). New York: Elsevier Inc.: Academic Press.
- Goodman, N. (1965). *Fact, fiction, and forecast*. Indianapolis, IN: Bobbs-Merrill Company.
- Heit, E. (2000). Properties of inductive reasoning. *Psychonomic Bulletin & Review*, 7, 569-592.
- Hernik, M., & Csibra, G. (2015). Infants learn enduring functions of novel tools from action demonstrations. *Journal of Experimental Child Psychology*, 130, 176-192.
- Hollander, M. A., Gelman, S. A., & Raman, L. (2009). Generic language and judgments about category membership: Can generics highlight properties as central? *Language and Cognitive Processes*, 24, 481-505.
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Kovács, Á. M., Téglás, E., Gergely, G., & Csibra, G. (2016). Seeing behind the surface: communicative demonstration boosts category disambiguation in 12-month-olds. *Developmental Science*.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, 674-689.

- Legare, C. H., Gelman, S. A., & Wellman, H. M. (2010). Inconsistency with prior knowledge triggers children's causal explanatory reasoning. *Child Development, 81*, 929-944.
- Legare, C. H. (2012). Exploring explanation: Explaining inconsistent evidence informs exploratory, hypothesis-testing behavior in young children. *Child Development, 83*, 173-185.
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175-220.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International University Press.
- Prasada, S. (2000). Acquiring generic knowledge. *Trends in Cognitive Sciences, 4*(2), 66-72.
- Prasada, S., & Dillingham, E. M. (2006). Principled and statistical connections in common sense conception. *Cognition, 99*, 73-112.
- Schmidt, M. F., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young Children See a Single Action and Infer a Social Norm Promiscuous Normativity in 3-Year-Olds. *Psychological Science, 27*, 1360-1370.
- Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology, 43*, 1045.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, MA: Harvard University Press.



- Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, 348, 91-94.
- Stahl, A. E., & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Cognition*, 163, 1-14.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535-585.
- Vredenburg, C., Kushnir, T., & Casasola, M. (2015). Pedagogical cues encourage toddlers' transmission of recently demonstrated functions to unfamiliar adults. *Developmental Science*, 18, 645-654.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129-140.
- Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the USA*, 105, 13690-13695.

**Figure 1.** Materials used in Experiment 1. (A) Music box toy, demonstration object, and categorization objects. (B) Bottom view of categorization objects with barcodes visible.



**Figure 2.** Percentage of children sorting by barcode or object shape in Experiment 1.

