The emergence of the empirical stance:

Children's testing of counter-intuitive claims

* *In Press at Developmental Psychology*

Samuel Ronfard[1], Eva E. Chen[2], & Paul L. Harris[3]

[1] *Boston University*

[2] *The Hong Kong University of Science and Technology*

[3] *Harvard University*

Correspondence concerning this article should be addressed to Samuel Ronfard.

Boston University
64 Cummington Mall St., Office 155
Boston, MA 02215
Phone: 617 974 0215
Email: sronfard@bu.edu

Acknowledgements. We appreciate the school staff for their support, and we thank the parents and children for their participation. We would also like to thank Serena Fan for advice and support throughout this project. We also thank Sze Long Poon and Kong Yan Tong for help with data collection and Han Pearl Li and Emma McCarthy for help with coding. This study was funded by a Sinclair Kennedy Travelling Fellowship from Harvard University to Samuel Ronfard and by an International Doctoral Student Visitor Supplement Fellowship from The Hong Kong University of Science and Technology to Samuel Ronfard. Finally, we wish to thank Vikram Jaswal, Jonathan D. Lane, and Deborah Kelemen for comments on an earlier version of this manuscript.

Abstract:

Although children often believe an adult's claims, they may have opportunities to check these claims by gathering relevant empirical evidence themselves. Here, we examine whether children seize such opportunities, especially when the claim is counter-intuitive. Chinese preschool and elementary school children were presented with five different-sized Russian dolls and asked to indicate the heaviest doll. Almost all children selected the biggest doll. Half of the children then heard a false, counter-intuitive claim (i.e., smallest = heaviest). The remaining children heard a claim confirming their initial intuition (i.e., biggest = heaviest). Children in both age groups typically endorsed the experimenter's claim even when it was counter-intuitive. However, during the experimenter's subsequent absence, elementary school children explored the dolls more if they had received counter-intuitive rather than confirming testimony whereas preschool children rarely explored, no matter what testimony they had received. Thus, with increasing age, children seize opportunities to test counter-intuitive claims.

Keywords: Cognitive Development; Learning; Problem Solving; Reasoning; Science Education; Testimony

The emergence of the empirical stance: Children's testing of counter-intuitive claims

To learn about the distant past, about remote places, or about hidden causal processes, children must typically rely on others' testimony (Harris & Koenig, 2006). A considerable body of evidence has shown that children are ready to trust what they are told in these various domains (Harris, 2012). Indeed, under certain circumstances, children are prone to trust testimony even when it contradicts their intuitions (Lane & Harris, 2014). Given that children often cannot gather direct, perceptual evidence to check what adults have told them – for example, they cannot easily gather evidence about the existence of germs or God – it is reasonable in such cases for children to accept what they have been told. However, there are situations where, in principle, children could test an adult's counter-intuitive claim. For example, if presented with similarly-sized cubes and told that some will float and others will sink, children could easily test the adult's claim and learn about the role of density through observation. Presented with a counter-intuitive claim that is easy to test via observation, do young children seize such opportunities or do they simply acquiesce to what they have been told?

One hypothesis is that young children will seek empirical evidence following such counter-intuitive claims. After all, infants seek such evidence after they have just seen a counter-intuitive phenomenon, such as a solid object moving through a wall (Stahl & Feigenson, 2015; see also Baldwin, Markman, & Melartin, 1993). Similarly, preschoolers seek evidence when faced with confounded (Schulz & Bonawitz, 2007) or theory-violating evidence (Bonawitz, Van Schijndel, Friel, & Schulz, 2012; Van Schijndel, Visser, Van Bers, & Raijmakers, 2015). However, an alternative hypothesis is that older – but not younger – children will seek evidence following a counter-intuitive claim. This hypothesis is plausible because younger children's response to *hearing* a counter-intuitive claim may be different from their response to *observing* a counter-intuitive phenomenon. When assessing testimony, children consider the reliability of the informant and the content of the testimony (Jaswal & Malone, 2007; Harris, Koenig, Corriveau, Jaswal, in press; Koenig & Echols, 2003). Younger children tend to place greater trust in adult testimony than older children. As a result, they may be less likely to investigate a surprising claim than older children.

Indeed, two- and three-year-old children display a robust bias to trust what other people tell them (Jaswal et al., 2014; Ma & Ganea, 2010). They are willing to trust information from a previously inaccurate or misleading adult (Krogh-Jespersen & Echols, 2012; Vanderbilt, Heyman, & Liu, 2014) – even one who has misled them multiple times (Jaswal, Croft, Setia, & Cole, 2010), and even when what they are told conflicts with what they have just seen (Jaswal, 2010). Such "blind" trust in testimony weakens over time (Ma & Ganea, 2010). Nevertheless, testimony continues to have a powerful effect on children's experience-based beliefs (c.f., the extensive literature on children's eyewitness testimony: Ceci & Bruck, 2006; Poole & Lindsay, 2001). Finally, younger preschoolers are more trusting of testimony than older preschoolers and elementary school-age children when learning about ambiguous entities and novel artifacts (Bernard, Harris, Terrier, & Clément, 2015; Lucas et al., 2016). Thus, younger children might not seek out evidence following a counter-intuitive claim, even if older children do so.

To assess these two hypotheses, we presented Chinese preschool and elementary school children with five Russian nesting dolls for visual inspection. The dolls varied in both size and weight – with the largest doll being the heaviest and the smallest doll being the lightest. We first asked children to say which doll was the heaviest. We expected almost all children to infer the dolls' weights based on their size (i.e., biggest = heaviest), because even infants as young as 9 months expect bigger objects to be heavier (Mounoud & Bower, 1974). Next, the experimenter made an assertion that either confirmed children's initial belief (i.e., biggest = heaviest) or contradicted it (i.e., smallest = heaviest, a false claim). To assess children's initial acceptance of the experimenter's claim, they were then re-questioned about the dolls' relative weight. Next, the experimenter excused herself from the room, thereby giving children an opportunity to explore the fit between their judgment about which doll was heaviest and the experimenter's claim – by lifting the dolls. Through comparing exploration of the dolls by children who had their initial judgment *confirmed* with exploration by children who had their initial judgment *contradicted*, we could assess the impact of confirming versus counter-intuitive testimony on children's exploration. To assess the possibility that children might restrict their exploration of the dolls because they assumed that they were not allowed to explore them, we randomly assigned approximately half the children within each testimony condition to receive a prime to explore the dolls (i.e., just before leaving, the experimenter pushed them within easy reach of the child).

When the experimenter returned, she first gave children an opportunity to comment on any discrepancy between her claim and what they had found via exploration of the dolls. The experimenter (E1) then asked children again about the relative weight of the dolls and left the room. Soon afterwards, a second experimenter (E2) entered the room, expressing interest in the dolls and noting that she had not seen them before. She then asked children about their weight, both directly (i.e., "Which one do you think is the heaviest?") and indirectly (i.e., "Which would make the best paperweight?"). Thus, we assessed the influence of children's exploration of the dolls on their subsequent judgments, not only when they were questioned by the adult who had informed them about their weights but also by an adult who seemed uninformed about them.

Method

Participants

We recruited a total of 190 children from one preschool and one elementary school in the city of Shenzhen, located in the Guangdong province of the People's Republic of China (92 female; M_{age} = 6.10 years, *Range* = 3.25 – 8.00 years). The two schools serve a similar population (described below) and are located in the same neighborhood (across the street from one another). We randomly assigned preschool and elementary school children to two conditions: counter-intuitive testimony and confirming testimony (see Table 1 for descriptive statistics). An a priori power analysis informed by a pilot study determined that we needed a minimum of 37 children in each Age Group X Testimony condition to have .8 power to detect a 35% difference in children's exploration across the two testimony conditions in each age group. Because the number of children who consented to participate exceeded expectations, we tested more children than originally planned.

We obtained a sample that was relatively diverse in family background. Parents reported on the level of education they and their partner had completed (186 out of 190, or 98% of parents answered this question) and on their income level (168 out of 190, or 88% of parents answered this question). Parents either: received no exposure to college (11%; i.e., neither parent had attended college), had some exposure to college (19%; i.e., at least one parent had attended college), or completed college (70%; i.e., at least one parent had completed college). Parents also reported having either: a higher-income level (7%), a middle-income level (87%), or a lower-income level (6%). The surveys were completed by children's mothers (77%), fathers (22%), or unspecified caregivers (<1%). Thirty-one additional children were recruited but not included in our analyses because of equipment failure (n = 20) or because children failed to identify the largest doll as the heaviest in their initial judgment (n = 11). An additional 10 elementary school children were recruited but not included because they were younger than the oldest preschool children. Not including these children ensured no overlap between age and school and allowed us to treat schooling (Preschool vs. Elementary School) as our age variable. Our statistical conclusions are robust to the inclusion of these children as well as to the exclusion of the 10 oldest preschoolers (rather than the 10 youngest elementary school children). This study – Children's Trust in Testimony – was approved by the Ethics Committees of Harvard University and of the Hong Kong University of Science and Technology (IRB#12421, HPR#121, respectively). Guardians of participants gave informed consent in writing before children participated in the study.

We recruited a sample of Chinese preschool and elementary school children for two reasons. First, testing Chinese children provided a potentially conservative estimate of children's testing of counter-intuitive claims. Prior research has found that Chinese and Chinese-American children tend to be more deferential to an adult's counter-perceptual testimony relative to American children (Corriveau, Kim, Song, & Harris, 2013; Corriveau & Harris, 2010; but see Chan & Tardiff, 2013). However, given that such differences only occurred when children were tested in a public rather than a private setting, we anticipated that our observation of children's independent testing of a counter-intuitive claim would be generalizable to other cultures. Our second reason for conducting this research with Chinese children is that most developmental research is conducted on American children, with very little conducted on non-WEIRD (Western, Educated, Industrial, Rich, and Democratic) samples (Henrich, Heine & Norenzayan, 2010). For example, Nielsen, Haun, Kärtner, and Legare (2017) found that 57% of participants in all articles published in leading developmental journals between 2006 and 2010 were from the United States, whereas only 4.36% were from Asian countries. This bias in the sampling of children in developmental studies may provide a limited picture of children's development. We judged that our testing of a non-U.S. sample could help to extend our understanding of children's cognitive development and provide an important point of comparison for future studies with samples from the United States and Europe.

Table 1.

Sample descriptive statistics by condition.

		Counter-Intuitive Testimony	**Confirming Testimony**
Preschool	*Prime*	*N* = 17, 9 female M_{age} = 4.75, *SD* = 1.06 *Age Range* = 3.42 to 6.25	*N* = 22, 9 female M_{age} = 4.74, *SD* = .89 *Age Range* = 3.33 to 6.25
	No Prime	*N* = 21, 9 female M_{age} = 4.47, *SD* = .91 *Age Range* = 3.25 to 6.42	*N* = 21, 13 female M_{age} = 4.58, *SD* = 1.00 *Age Range* = 3.33 to 6.33
Elementary School	*Prime*	*N* = 27, 14 female M_{age} = 7.21, *SD* = .40 *Age Range* = 6.48 to 7.87	*N* = 28, 16 female M_{age} = 7.19, *SD* = .36 *Age Range* = 6.62 to 8.00
	No Prime	*N* = 26, 11 female M_{age} = 7.22, *SD* = .38 *Age Range* = 6.48 to 7.86	*N* = 28, 11 female M_{age} = 7.14, *SD* = .32 *Age Range* = 6.62 to 7.90

Materials

We used five different-sized Russian nesting dolls; each doll was attached to a square base for stability. With the square base attached, the dolls weighed (from smallest to biggest): 16.32 g, 29.04 g, 46.75 g, 85.82 g, and 167.73 g. Thus, size and weight were correlated – the smallest doll was also the lightest and the biggest doll was the heaviest. The dolls and their bases were painted white. They were arranged on a tray placed on the table so that the biggest doll was on the child's left and the smallest doll was on the child's right (Figure 1). The experimenter and the child sat next to each other at the table. The dolls were approximately 18" from the table edge nearest to the child. The experimental session was discreetly recorded using a laptop camera with a darkened screen. None of the children made any comments about the laptop or behaved as if they knew they were being filmed.

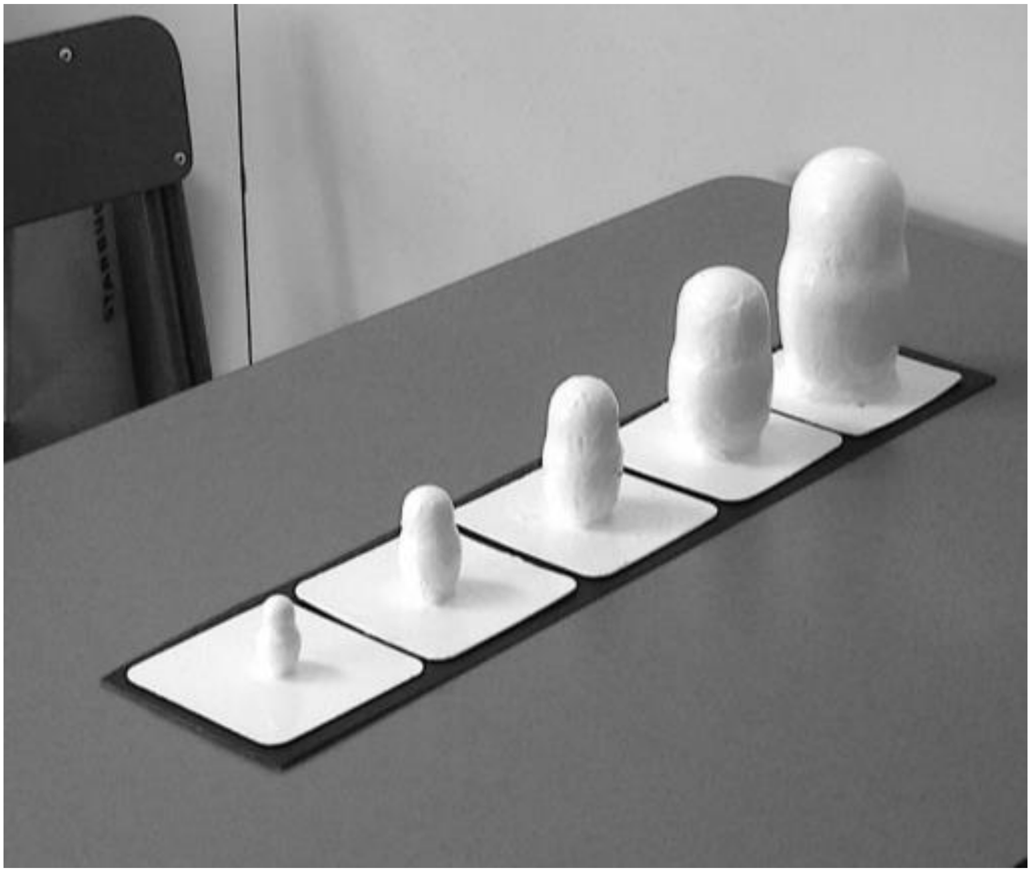

Figure 1. Stimuli used in the experiment

Procedure

Children were individually tested in a separate room at their school by a female Chinese experimenter fluent in Mandarin. The experimental procedure consisted of eight phases: (i) initial judgment with E1; (ii) testimony (counter-intuitive or confirming, depending on condition); (iii) post-testimony judgment; (iv) opportunity to explore the dolls (prime or no prime, depending on condition); (v) opportunity to report exploration to E1; (vi) final judgment with E1; (vii) initial judgment with E2; and (viii) paperweight selection with E2. A brief description of each phase follows.

Initial Judgment with E1. The experimenter asked children to point to the heaviest doll: "Which doll do you think is the heaviest?" Children were then asked an open-ended question inviting them to explain their judgment: "You think this one is the heaviest – why do you think it is the heaviest?"

Testimony. Children were randomly assigned to receive either counter-intuitive testimony (i.e., smallest = heaviest) or testimony that confirmed their intuition (i.e., biggest = heaviest). In the counter-intuitive testimony condition, the experimenter told children: "Actually, that one is not the heaviest; this one here (pointing to the smallest one on the right) is the heaviest. It's heavier than all of the other ones. It's heavier than this one, this one, this one, this one (starting with the biggest one and moving to the second smallest one)." Note that this statement was false because the smallest dolls was the lightest and the biggest doll was the heaviest. In the confirming testimony condition, the experimenter told children: "Yes, that one is the heaviest, and this one here (pointing to the smallest one on the right) is the lightest. This one (pointing to the biggest one) is heavier than all of the other ones. It's heavier than this one, this one, this one, and this one (starting with the second largest one and moving to the smallest one)".

Post-testimony judgment. Children were again asked to identify the heaviest doll and to provide an explanation for their judgment using the same wording as for the initial judgment: "Which doll do you think is the heaviest?" Children were also asked to recall which doll the experimenter had identified as the heaviest: "Can you point to the one I said was the heaviest?" Of the 190 children tested, 176 (93%) correctly pointed to the doll indicated by the experimenter. The remaining 14 children (10 in the counter-intuitive testimony condition and 4 in the confirming testimony condition) selected a medium-sized doll (i.e., a doll that was not the smallest doll but that was not the biggest doll either).

Opportunity to explore the dolls. The experimenter then told children that she was going next door to use her phone for a moment but that she would come right back. For children assigned to the prime condition, she added: “I’ll move the dolls a bit closer to you” and pushed the tray so that the dolls were about 6 inches from the child. She then walked out of the room, returning, on average, after 76 seconds had elapsed (*SD* = 6 seconds).

Opportunity to report exploration to E1. Once the experimenter returned, she said, “Let’s see – we were talking about the dolls,” and paused for 10 seconds to offer children an opportunity to initiate a conversation with her following their opportunity to explore the dolls. If children did not spontaneously talk to her, she prompted children: “Okay, we’re almost done. Is there anything you want to tell me?”

Final judgment with E1. Children were asked to identify the heaviest doll and to provide an explanation for their judgment using the same wording as the initial judgment. E1 then stated: “We’re almost done, please wait in this room and somebody will come to take you back to the classroom.” She left the room and went to get E2, a different female experimenter. She did not return to the room but E2 entered the room. Note that the child had not yet met E2 who had remained in another room for the first part of the experiment.

Initial judgment with E2. E2 entered the room and exclaimed: “Whoa! I like these dolls. I've never played with them before! I wonder which one is the heaviest? Can you tell me?” E2 then asked for an explanation by saying: “Why do you think that one is the heaviest?”

Paperweight selection with E2. E2 then stated: “Hey—these dolls give me an idea! I think one of these dolls would be good to stop my papers from blowing away—especially if it’s heavy (E2 puts a pile of papers on the table). Can you point to the doll you think is best to stop the papers from blowing away?" After the child made a choice, E2 asked for an explanation: "Why do you think that one is best to stop the papers from blowing away?"

Coding

The first author and a research assistant blind to the hypotheses of the study coded 100% of the videos for children's exploration of the doll. Both coders were blind to children's age, condition, and judgments about the dolls. The inter-rater correlation for the number of dolls children picked up was $\rho = .994$. Discrepancies were resolved through discussion. Our analyses are based on the data coded by the research assistant blind to the hypotheses of the study.

Results

We analyze children's: (i) initial and post-testimony weight judgments; (ii) exploration of the dolls; (iii) post-exploration remarks to E1; (iv) post-exploration weight judgments; (v) and post-exploration weight judgments as a function of children's exploration. We report results for logistic regression models using odds-ratios. All logistic analyses were conducted using the –logit– command in Stata 14.

Children's initial and post-testimony weight judgments

As Table 2 shows, and by design, all children initially stated that the biggest doll was the heaviest (as noted above, we excluded 11 children [5.8% of the sample] who did not do so). When asked to make the post-testimony judgment, almost all children who heard testimony that confirmed their initial judgment continued to make the same judgment, McNemar tests > .25. In contrast, very few of the children who heard testimony that conflicted with their initial judgment continued to make the same judgment, McNemar tests < .0001. They endorsed the biggest doll as the heaviest significantly less often than children who received confirming testimony, $\chi^2(1, N = 190) = 67.50$, $p < .001$, Cramér's $V = .60$. In sum, the type of testimony markedly affected children's judgments of the doll's weights. In Online Supplemental Material -S1, we report children's explanations following their initial judgment about the dolls' weight and following their subsequent judgment of the dolls' weight after receiving testimony.

Table 2. Percentage of children who claimed that the biggest doll was the heaviest.

		Initial Judgment	**Following Testimony**
Confirming	*Preschool*	100%	93%
	Elementary	100%	96%
Counter	*Preschool*	100%	13%
	Elementary	100%	8%

Children's exploration of the dolls

We conducted two rounds of coding. In an initial 'generic' coding, we coded whether or not children touched any of the dolls (irrespective of size) when the experimenter was out of the room. This allowed us to assess whether the receipt of counter-intuitive testimony increased children's general interest in the dolls (i.e., touching the dolls). We analyzed children's touching of the dolls using a logistic regression model to assess the influence of Age Group (Elementary vs. Preschool), Testimony Type (Counter-Intuitive vs. Confirming), and Priming (Prime vs. no Prime) on the probability that children touched the dolls. We then tested for interaction effects.

Table 3 displays parameter estimates for a logistic regression model predicting whether children touched any of the dolls. All coefficients are in odds-ratios. For example, the coefficient for Counter-Intuitive Testimony is the ratio of the odds that a child touched any of the dolls after having received counter-intuitive rather than confirming testimony.

There was a main effect of Age Group, indicating that elementary school children were more likely to touch the dolls than preschoolers, as displayed in Figure 2. There was no main effect of Testimony Type. Thus, at both ages, a similar proportion of children engaged in general exploration of the dolls via touching, regardless of whether they received confirming or counter-intuitive testimony. Note that most children in each condition and age group touched the dolls. There was no main effect of Priming. Thus, whether children received a prime to explore or did not receive a prime to explore, a similar proportion of children touched the dolls. The various interactions between Age Group, Testimony Type, and Priming were not statistically significant.

Table 3. Logistic regression model comparing whether children touched the dolls as a function of the type of testimony children received, whether they received a prime, and their age.

	Odds-Ratios	**z scores**	**95% CI**
Counter-Intuitive Testimony	1.28	.69	.64, 2.58
Elementary	3.89***	3.75	1.91, 7.92
Priming	1.57	1.25	.78, 3.16
Constant	1.16	.46	.61, 2.22
X^2		17.22***	
Model df		3	
-2 Log Likelihood		193.11	

*** $p < .001$. *Note*. $n = 190$.

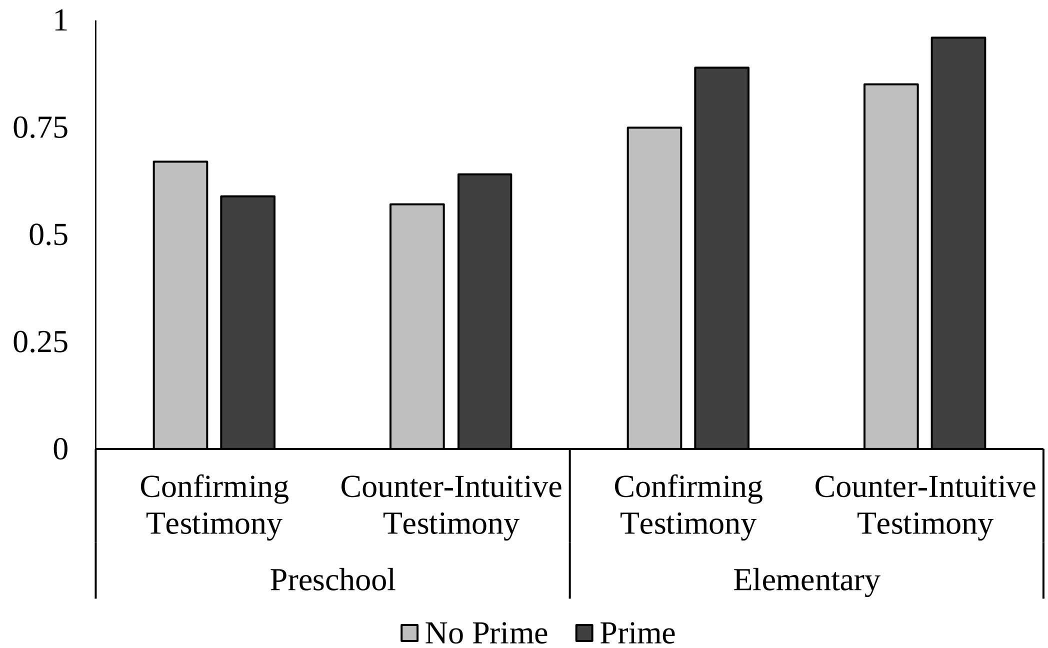

Figure 2. Proportion of preschool and elementary school children who touched the dolls while E1 was out of the room as a function of the testimony and priming they received prior to E1's departure.

To investigate children's exploration in more detail, we coded in 5-second intervals whether or not children picked up a doll, which particular doll they picked up, and how many times they picked up that doll. In Figure 3, we display the proportion of preschool and elementary school children who picked up each doll when the experimenter left the room (coded in successive 5-second intervals) for the confirming and counter-intuitive testimony conditions. Inspection of Figure 3 reveals that throughout the experimenter's absence, preschool children did not pick up the dolls differentially whether they had received counter-intuitive or confirming testimony. Thus, irrespective of the testimony they had heard, a small proportion of preschoolers picked up each doll throughout the absence of E1. In contrast, there was a clear difference in the proportion of elementary children who picked up the smallest and the biggest doll in the counter-intuitive and confirming testimony conditions. This difference was especially apparent immediately after the experimenter left the room and again after an additional 50 seconds had elapsed. This suggests that elementary school, but not preschool, children sought evidence that could confirm or disconfirm the experimenter's counter-intuitive testimony.

To confirm this interpretation, we calculated the number of times each child picked up each doll during the experimenter's absence (Figure 4) and conducted a repeated-measures ANOVA with the between-subject factors of Age Group (2: Elementary vs. Preschool), Testimony Type (2: Counter-Intuitive vs. Confirming), and Priming (2: Prime vs. no Prime), and the within-subject factor of Doll (5: one [i.e., smallest], two, three, four, and five [i.e., biggest]) on the number of times children picked up a doll. This analysis revealed a modest effect of Priming, $F(1, 182) = 6.12$, $p < .05$, $\eta^2_p = .03$. Children picked up the dolls more often when they had received a prime to explore than when they had not, $M = 6.68$, $SD = 6.99$ vs. $M = 4.19$, $SD = 4.14$. However, receiving a prime to explore did not interact with Age Group, Testimony Type, or Dolls. By implication, receiving a prime to explore increased children's general exploration of the dolls (i.e., whether they picked up the dolls), but it did not increase their targeted exploration of the dolls (i.e., whether they picked up a particular doll).

Our analysis also revealed a main effect of Doll, $F(4, 728) = 29.35$, $p < .001$, $\eta^2_p = .14$, a main effect of Age Group, $F(1, 182) = 16.32$, $p < .001$, $\eta^2_p = .08$, a significant Age Group X Testimony Type interaction, $F(1, 182) = 5.98$, $p = .015$, $\eta^2_p = .03$, a significant Doll X Testimony Type interaction, $F(4, 728) = 7.10$, $p < .001$, $\eta^2_p = .04$, a significant Doll X Age Group interaction, $F(4, 728) = 8.72$, $p < .001$, $\eta^2_p = .05$, all of which were subsumed under the significant 3-way interaction of Doll X Age Group X Testimony Type, $F(4, 728) = 3.59$, $p =$

.007, $\eta^2_p = .02$. To better understand this three-way interaction, we computed simple tests of the effect of testimony for each of the ten combinations of age and doll using a Bonferroni correction. For preschool children, the frequency with which they picked each doll was not influenced by the type of testimony they received, $p > .13$, for all five tests. By contrast, elementary school children picked up the smallest and the biggest dolls significantly more often when they received counter-intuitive rather than confirming testimony, both $p < .001$. Testimony type did not influence the frequency with which elementary school children picked up the three intermediate dolls, $p > .14$ for all three tests. Thus, elementary school children, but not preschool children, sought evidence (lifting the biggest and smallest dolls) that could confirm or disconfirm the informant's testimony. Moreover, this targeted exploration cannot be attributed to a general tendency of elementary school children to pick up the dolls more often than preschool children. Preschool and elementary school children in the confirming testimony condition did not differ in the frequency with which they picked up all five dolls, $p > .20$ for all five dolls.

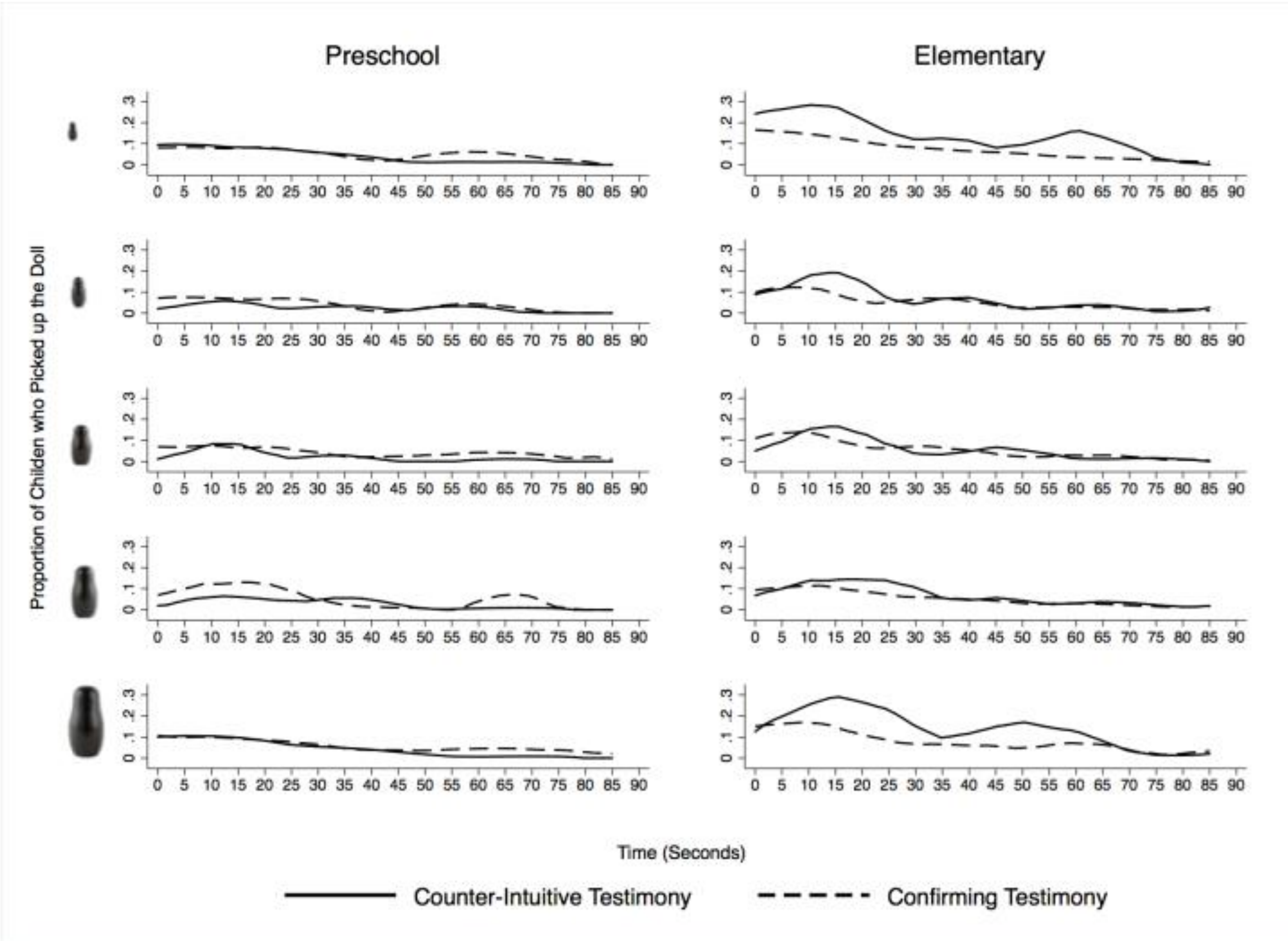

Figure 3. Proportion for children who picked each doll when the experimenter left the room (coded in 5-second intervals). Dolls are displayed in black for easier viewing.

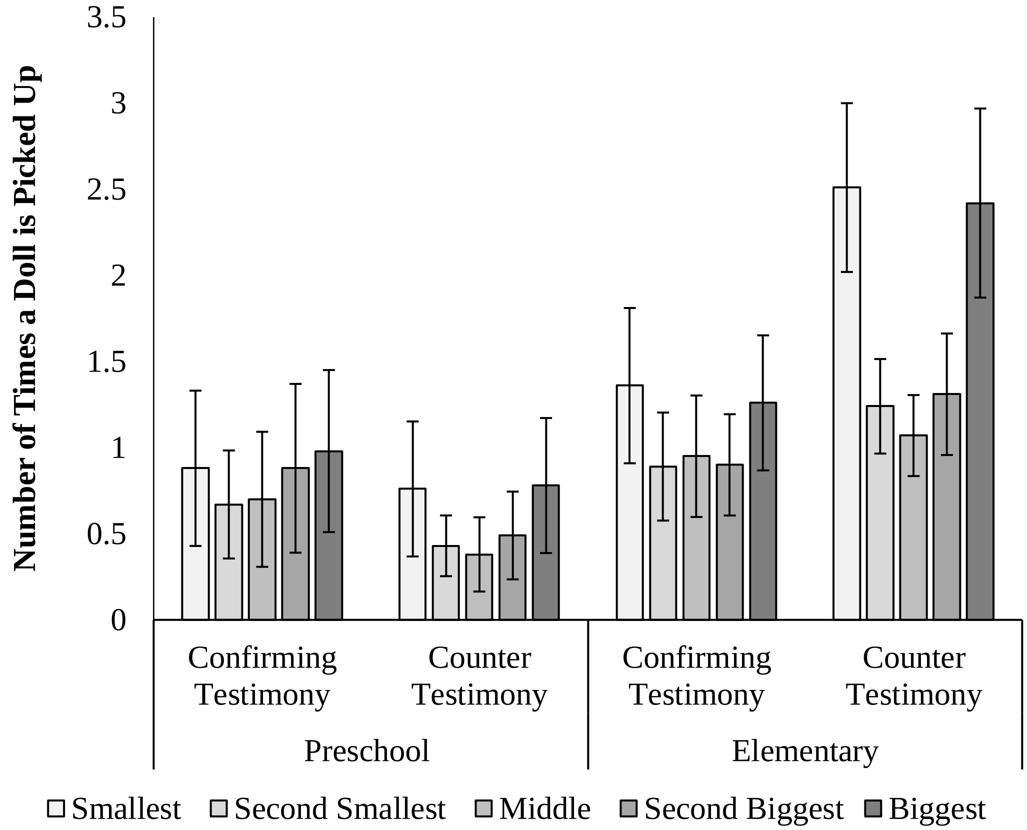

Figure 4. Number of times a doll was picked up when the experimenter left the room. Error bars represent 95% confidence intervals.

In summary, elementary school children's targeted exploration of the dolls (i.e., picking up the smallest and biggest dolls) but not their general interest in the dolls (i.e., simply touching any of the dolls) was influenced by the counter-intuitive nature of the testimony. By implication, counter-intuitive testimony provoked elementary school children to compare the biggest and smallest dolls. Indeed, they engaged in such exploration whether or not they were primed to do so by E1's pushing the dolls toward them prior to making her exit. This effect of counter-intuitive testimony was not observed among preschool children.

In Online Supplemental Material-S2, we replicate our analyses of children's exploration of the dolls (i.e., picking up the dolls) when we exclude the minority of children who did not touch any of the dolls (i.e., 38% of preschool children and 14% of elementary school children). In Online Supplemental Material-S3, we demonstrate the same pattern of results using a still more targeted measure of children's exploration – notably, whether they picked up the biggest and the smallest doll *at the same time,* thereby optimizing their opportunity to establish which doll was heavier. Also, in Online Supplemental Material-S4, we show that children's decision to pick up both the biggest and the smallest doll at some point during the experimenter's absence in the counter-intuitive testimony condition was unrelated to whether they had endorsed or rejected the experimenter's testimony and to the type of explanation they had provided following their endorsement or rejection of the experimenter's claim.

Post-exploration remarks to the informant

Following the return of the informant (E1), some children commented on the weight of the dolls (e.g., "The smallest doll is the lightest"; "I know which doll is the heaviest! [pointing to the largest one]. I have picked up all of the dolls"). Although rare, these comments on the weight of the dolls displayed a similar pattern to children's doll exploration. Whereas preschoolers seldom made them following either type of testimony (3 vs. 2, one-sided Binomial Test, $p > .25$) elementary school children made them more often following counter-intuitive as compared to confirming testimony (9 vs. 2, one-sided Binomial Test, $p = .033$). That said, the majority of children, irrespective of age and testimony, did not comment on the weight of the dolls.

Post-exploration weight judgments

Children made a judgment about the weight of the dolls: (i) when explicitly asked by E1 immediately after the opportunity to explore the dolls; (ii) when explicitly asked by E2; and (iii) when invited by E2 to select the heaviest paperweight. In Table 4, we display the proportion of children for each age group who endorsed the biggest doll as the heaviest at these three time points. Inspection of Table 4 and statistical analyses reveal that children's judgments remained relatively stable across these three successive time points (see Online Supplemental Material-S5). Accordingly, we summed the three judgments to reflect how often children asserted that biggest = heaviest following their opportunity to explore the dolls.

Table 4. Proportion of preschool and elementary school children who endorsed the biggest doll as the heaviest (1) immediately after the opportunity to explore the dolls when questioned by E1; (ii) when explicitly asked by E2; and (iii) when invited by E2 to select the heaviest paperweight.

	Following Opportunity to Explore with E1	**Initial Judgement with E2**	**Paperweight Task with E2**
Confirming			
Preschool	98%	93%	77%
Elementary	96%	100%	91%
Counter-Intuitive			
Preschool	34%	37%	45%
Elementary	49%	53%	68%

Children's combined scores were analyzed via an ANOVA with Age Group (2: Preschool, Elementary), Priming (2: Prime, No Prime), and Testimony Type (2: Confirming, Counter-Intuitive) as between-subject factors. This analysis revealed significant main effects of Testimony Type, $F(1,182) = 90.14$, $p < .001$, $\eta^2_p = .34$, and Age Group, $F(1,182) = 4.84$, $p = .026$, $\eta^2_p = .027$. There were no other significant main effects or interaction effects. Although the interaction of Testimony Type X Age Group was not significant, $F(1,182) = 2.14$, $p = .14$, $\eta^2 = .012$, we conducted Bonferroni corrected post-hoc tests of the simple effect of age for each type of testimony given the aforementioned age-related differences in children's sensitivity to the type of testimony they received, as reflected in their exploration of the dolls. Figure 5 shows the frequency with which children said that the biggest doll was the heaviest as a function of Age Group and Testimony Type. Inspection of Figure 5 reveals that preschool and elementary school children who received confirming testimony made similar judgments, $p < .25$. In contrast, in the counter-intuitive condition, elementary school children judged that the biggest doll was the heaviest more often than younger children, $p = .011$.

Note that this conclusion is further supported when looking at the data in Table 4. Elementary school and preschool children in the confirming testimony condition overwhelmingly endorsed the biggest doll as the heaviest at all three time points and did so at similar levels. In contrast, in the counter-intuitive testimony condition children were less likely to select the biggest doll at the heaviest across all three time points. Though this was true for preschool and elementary school children, elementary school children endorsed the biggest doll as the heaviest more often than preschool children at all three time points.

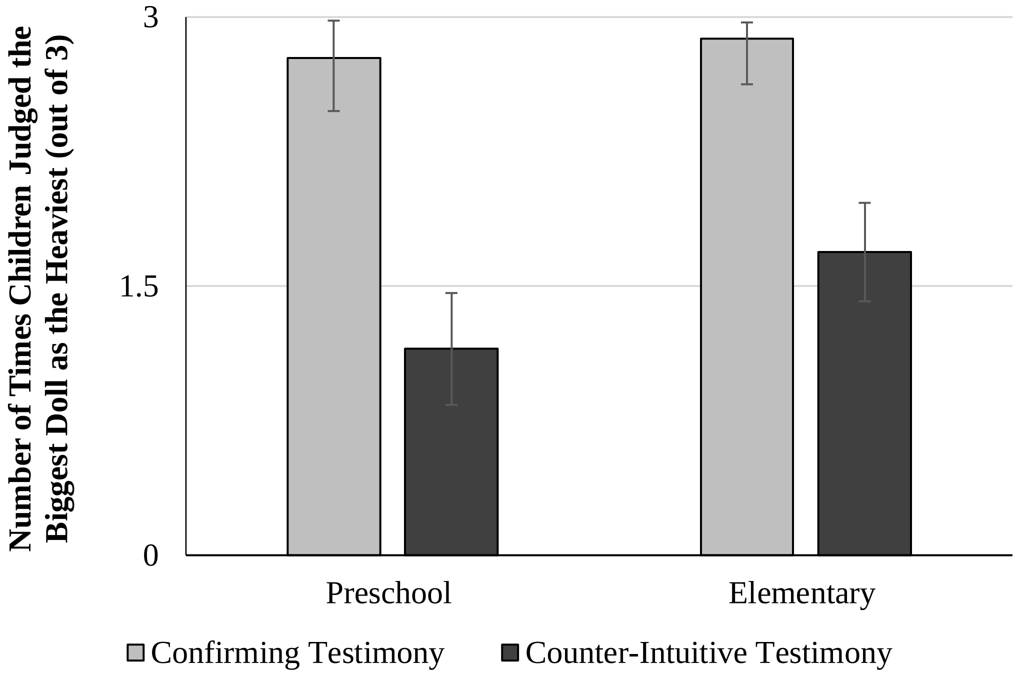

Figure 5. Average number of times preschool and elementary school children judged that the biggest doll was the heaviest as a function of the type of testimony they received. Error bars represent 95% confidence intervals.

Post-exploration weight judgments as a function of children's exploration

Finally, to assess whether children's exploration had impacted their subsequent weight judgments, we analyzed these judgments via an ANOVA with Age Group (2: Preschool, Elementary), Priming (2: Prime, No Prime), and Exploration (2: Explored, Did Not Explore) as between-subject factors, restricting our analysis to children who had received counter-intuitive testimony. Note that we operationalized exploration as children's decision to pick up the biggest *and* the smallest doll at some point during the experimenters' absence. However, in Online Supplemental Material-S6, we replicate the results we report below using a less stringent measures of exploration, i.e., whether children picked up the smallest *or* any other doll during the experimenter's absence. We conducted this additional test because children did not need to pickup the smallest and the biggest doll to falsify the experimenter's claim. They could have simply picked up the smallest doll and any other doll.

This analysis revealed only a significant main effect of Exploration, $F(1,83) = 21.79$, $p < .001$, $\eta^2_p = .21$: Children who had explored judged the biggest doll to be the heaviest much more often than children who had not explored. We display this main effect in Figure 6. Children who did not explore the dolls endorsed the smallest doll as the heaviest significantly above chance, $t(33) = 4.42$, $p < 0.001$, $d = 1.54$. In contrast, children who did explore the dolls, endorsed the biggest doll as the heaviest significantly above chance, $t(56) = 2.83$, $p = 0.006$, $d = .75$. Thus, when children gathered empirical evidence, it undermined the earlier impact of E1's counter-intuitive testimony on their judgments. The absence of any interaction between age and exploration, $F(1,83) = .30$, $p > .25$, indicates that when preschool or elementary school children explored, it impacted their judgments to the same extent, as shown in Figure 6.

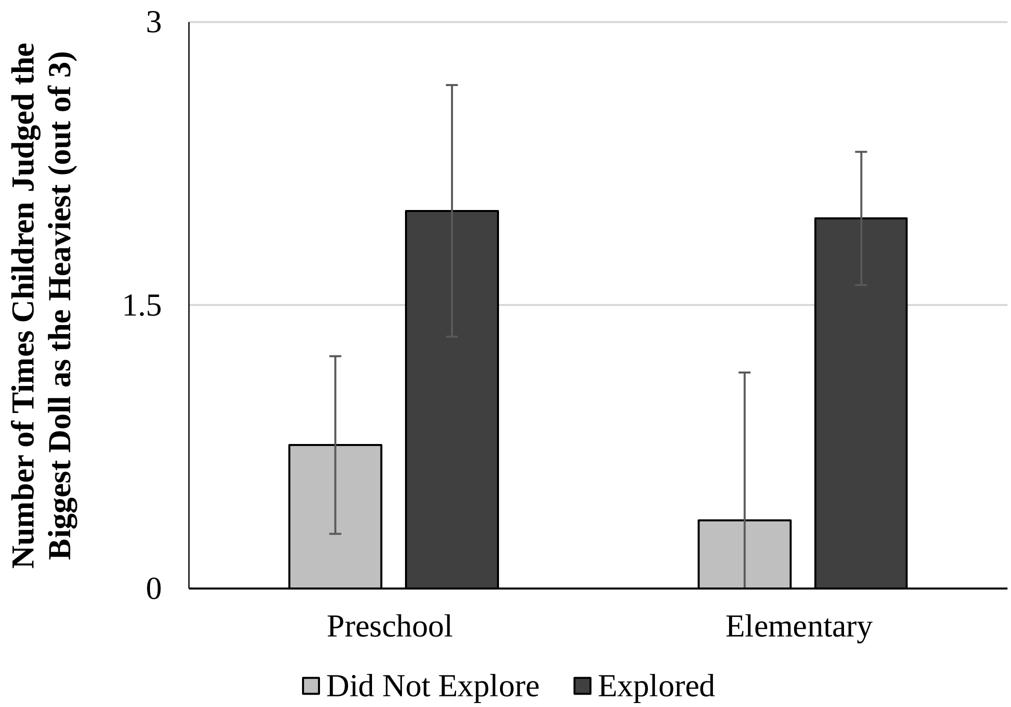

Figure 6. Average number of times preschool and elementary children judged that the biggest doll was the heaviest as a function of whether they picked up the biggest and the smallest dolls (Explored: 44 elementary school children, 13 preschool children; Did Not Explore: 9 elementary school children, 25 preschool children). Error bars represent 95% confidence intervals.

Discussion

We asked whether children would seek empirical evidence following an experimenter's testable claim. Preschool children rarely did so whether the claim matched their intuitions or conflicted with them. By contrast, elementary school children did seek empirical evidence, especially when the claim was counter-intuitive. We first discuss the pattern of results for elementary school children and the extent to which they cohere before turning to those obtained for preschool children.

The experimenter's counter-intuitive claim triggered elementary school children's search for evidence confirming or disconfirming that claim, which led them to revise their initial endorsement of the experimenter's testimony. Several key findings support this claim. First, we found that it was elementary school children's targeted exploration of the dolls (i.e., picking up the biggest and the smallest doll) and not their more general interest in the dolls (i.e., simply touching the dolls) that was influenced by the counter-intuitive nature of the testimony. Indeed, elementary school children picked up the smallest and the biggest dolls more often (separately and concurrently) when they had received counter-intuitive rather than confirming testimony. Moreover, analyses of children's comments to the experimenter upon her return to the room showed that elementary school children not only sought empirical evidence more often following a counter-intuitive claim, but also queried the experimenter more often about the weight of the dolls. Children's empirical investigations of the dolls also had a marked effect on their judgments about their weight. Following initial receipt of counter-intuitive testimony, a clear majority of elementary school children endorsed the first experimenter's claim that the smallest doll was the heaviest (92%). However, following the opportunity to explore the dolls, children endorsed her claim only 44% of the time (i.e., 1.31 out of 3 times). Importantly, this reversal in children's judgments was tied to their actual exploration of the dolls. When children explored the dolls, they reneged on their endorsement of the experimenter's testimony and endorsed the biggest doll as the heaviest 65% of the time (i.e., 1.96 out of 3 times). In contrast, when children did not explore the dolls, they endorsed the biggest doll as the heaviest 12% of the time (i.e., 0.36 out of 3 times). In sum, we found that elementary school children are prone to adopt an empirical stance in relation to counter-intuitive claims. Recognizing that such claims can be tested against the available evidence, they actively seek that evidence and revise their judgment of the unexpected claim accordingly. Thus, elementary school children engage in selective exploration not just following their observation of a counter-intuitive event (Bonawitz, et al., 2012; Sodian, Zaitchik, & Carey, 1991; Van Schijndel et al., 2015), but also following their receipt of a counter-intuitive claim.

In contrast, there was limited evidence that preschool children adopt an empirical stance in relation to the experimenter's counter-intuitive claim. Like the elementary school children, they began the experiment stating that the biggest doll was the heaviest and, by implication, were surprised by the experimenter's claim that the smallest doll was the heaviest. However, we found no indication of any selective exploration by preschoolers following counter-intuitive testimony as compared to confirming testimony. Thus, there was no difference in preschool children's tendency to touch the dolls across the two testimony conditions, and we found no significant differences in the frequency with which preschool children picked up any of the five dolls across conditions – a result that remained consistent even when we excluded the minority of children who never touched any of the dolls. If preschoolers did explore the dolls – and a minority did so – they responded in the same way as elementary school children. They used the evidence they gathered to re-assess the experimenter's claim and they reverted to their initial (and correct) belief about the relationship between size and weight.

It is worth noting that three explanations for this age change received no support. First, there was no indication that the two age groups varied in their ability to learn from empirical evidence, once it had been gathered. Recall that children in each age group were equally likely to reaffirm – in the wake of exploration – that the biggest doll was the heaviest (See Figure 6). By implication, the two age groups did not differ in their capacity to revise their judgment once empirical evidence had been gathered. Instead, they differed in their readiness to seek such evidence in the first place, especially after counter-intuitive testimony (see Figure 4).

Second, there was no indication that preschool children explored the dolls less than elementary school children because they believed that such exploration was unacceptable. Receiving a prime to explore the dolls did not have a differential impact on younger and older children's willingness to touch the dolls or on the number of dolls that they picked up. Moreover, preschool and elementary school children's selective exploration of the dolls differed significantly from each other only in the counter-intuitive testimony condition. Indeed, in the confirming testimony condition, preschool and elementary school children did not differ in the number of times they picked each of the five dolls. By implication, it is not the case that older children were generally more bold or exploratory than younger children.

A third explanation of the age change also receives no clear support. Children's responses to the counter-intuitive testimony might reflect age-related differences in their initial intuitions about the dolls or age-related difference in their confidence in those initial intuitions. However, there was no indication that the two age groups differed in their initial intuitions. Recall that almost all children in both age groups claimed that the biggest doll was the heaviest. By implication, children in both age groups found the experimenter's counter-intuitive claim unexpected. In addition, there was no indication that the two age groups differed in the confidence that they had in their initial intuitions because a similar proportion of children in each age group endorsed the experimenter's counter-intuitive testimony.

Admittedly, there was an age change in the way that children explained their endorsement of the experimenter's counter-intuitive claim – with a considerable proportion of older children noting that size is sometimes unrelated to weight, a finding consistent with the earlier findings of Smith, Carey, and Wiser (1985). However, the content of children's explanations proved to be unrelated to their pattern of exploration. In Online Supplemental

Material-S4, we show that children's decision to pick up both the biggest and the smallest doll during the experimenter's absence in the counter-intuitive testimony condition was unrelated to the type of explanation they had provided following their endorsement or rejection of the experimenter's claim. Moreover, although children's explanations for their endorsement of the experimenter's testimony point to age changes in their understanding of the conceptual distinction between weight and mass, it is plausible that these conceptual differences would make younger children more surprised than older children about the experimenter's testimony because older children would have found it easier to represent and hence accept the experimenter's counter-intuitive claim that the smallest doll was the heaviest. Indeed, prior work shows that greater conceptual understanding can lead to increase trust in counter-intuitive claims. For example, Lane, Harris, Gelman, and Wellman (2014) found that preschoolers who had a greater understanding of the appearance-reality distinction were more likely to accept an experimenter's counter-perceptual claim (e.g., that an object that looked like a rock was in fact soap). Taken together, these findings imply that older children did not explore more than younger children because they found the experimenter's claim more unexpected than did younger children.

Two more plausible explanations of the age change focus on age-related differences in children's stance toward adult testimony (i.e., differences in beliefs about the possibility that an adult would provide misleading testimony) and on age-related differences in children's realization that they could pick up the dolls to establish the truth or falsity of the experimenter's claim. We consider each of these two possibilities in turn.

First, as discussed in the introduction, younger children tend to be more trusting of adult testimony than older children. This greater trust in adult testimony may have led younger children to feel less uncertainty about the experimenter's counter-intuitive claim than older children, thereby reducing the need for exploration. When assessing a counter-intuitive claim, children consider the fit between the claim and the physical evidence and the reliability of the informant (Bridgers et al., 2016; Harris et al., in press; Sobel & Kushnir, 2013). Our analyses suggest that younger and older children had relatively similar intuitions about the physical evidence in front them. By implication, older children's greater investigation of the dolls following counter-intuitive testimony may reflect greater skepticism towards adult testimony –a claim consistent with previously reported age-related decreases in counter-perceptual testimony provided by adult informants (e.g., Bernard et al., 2015) and with general developments in children's theory of mind, particularly in their appreciation that speakers may intentionally make claims they do not believe (e.g., children's understanding of irony, cynicism, and lying, Filippova & Astington, 2008; Mills & Kiel, 2005; Peterson et al., 2012; Talway & Lee, 2008; Wimmer, Gruber, Perner, 1984).

A second explanation for the age-related difference in children's exploration of the dolls is that younger and older children differed in their realization that they could pick up the dolls to establish the truth or falsity of the experimenter's claim. Indeed, there is evidence that children's reasoning about how to gather information improves considerably during the preschool and elementary school years. For example, Fitneva, Lam, and Dunfield (2013) asked 4- to 6-year-old children to choose whether to engage in visual inspection or to ask an expert about visible and invisible properties of entities (e.g., color vs. knowledge of French). It was not until children were 6 years of age that they were above chance in asking the expert about invisible properties and engaging in visual inspection for visible properties. Similarly, in a longitudinal study that tested children at 4-, 5-, and 6-years-old, Piekni, Grube, and Maehler (2014) found a significant increase in children's ability to select an appropriate experimental test between 5- and 6-years-old (see also Croker & Buchanan, 2011). The authors told children that two brothers disagreed about the size of a mouse in their basement – one brother thought it was big and the other one thought it was small. They then asked children two questions. First, they asked children whether the brothers should cut out a small door or a big door in a box that contained food if the brothers wanted to feed the mouse (correct answer = a big door). Second, they asked children whether the brothers should cut out a small door or a big door in that box if they wanted to know the size of the mouse (correct answer = a small door) (i.e., a task similar to that used by Sodian, Zaitchik & Carey [1991] with first and second graders). Piekni and colleagues (2014) found that only 31% of 4- and 5-year-old children answered these two questions correctly, whereas almost 47% of 6-year-olds did so. These results echo our own. It was not until elementary school (i.e., 6.5- to 8-years old) that children displayed a clear recognition that picking up the dolls would help them to resolve the conflict between their initial intuition and the testimony they were given. Admittedly, both Fitneva and colleagues (2013) and Piekni and colleagues (2014) focused on children's ability to choose between empirical tests, whereas we focused on whether children spontaneously engage in empirical testing. However, children's ability to reason about how to test a claim is likely to impact their ability to test it – particularly when the claim involves two dimensions (i.e., height and weight). Given that our two age groups differ in the length and type of schooling (i.e., preschool vs. elementary school) they have received, it is also possible that the age-related differences in exploratory testing that we found reflect differences in children's exposure to formal instruction about hypothesis testing and the search for confirming evidence – a skill that is particularly difficult to acquire without direct instruction (Chen & Klahr, 1999).

In Online Supplemental Material –S7, we conducted post-hoc exploratory analyses to test these two conflicting hypotheses: cognitive maturation vs. schooling. We tested for a significant interaction between age (in months) and the kind of testimony children received within each school group (i.e., preschool and elementary school). If there is an effect of age (within school group), we would expect a positive interaction between age and whether children received counter-intuitive testimony, i.e., older children are more likely to explore the dolls in the counter-intuitive testimony condition than in the confirming testimony condition. Our analyses do not find such an interaction. This is consistent with the claim that instruction at the elementary school level explains developments in children's empirical stance. However, these post-hoc tests do not rule out the cognitive maturation hypothesis. More research comparing children who begin elementary school at different ages is needed to more fully test these two hypotheses. Overall, however, our results provide strong evidence that by the start of elementary school, children seize opportunities to gather empirical evidence to test counter-intuitive claims.

We consider three implications of the main finding – the targeted exploration of elementary school children when given counter-intuitive testimony. This finding adds an important twist to prior research on children's reactions to conflicts between what they see and what they are told. As outlined in the introduction, this line of research has shown that when presented with testimonial evidence conflicting with their perceptions and intuitions, children often resolve that conflict by deferring to the testimony (e.g., Gelman & Markman, 1986; 1987). Indeed, children will even pass on such counter-perceptual claims to someone else (Jaswal, Lima, & Small, 2009). Our results show that although children may be willing to entertain, endorse, and even transmit counter-intuitive claims, their endorsement of such claims need not imply unreflective acceptance of them. Particularly among elementary school children, the counter-intuitive claim prompted targeted empirical exploration.

Second, this finding extends prior work on the impact of instruction on children's exploration. Previous research has shown that, under certain conditions, instruction restricts exploration by reducing the number of hypotheses that children consider. When a toy's functions are demonstrated in a pedagogical rather than a non-pedagogical context, children infer that only the demonstrated functions are relevant and focus their exploration on those functions rather than on discovering additional functions (e.g., Bonawitz et al., 2011; Shafto, Goodman, & Frank 2012; Shneidman, Gweon, Schulz, & Woodward, 2016). Our results extend this work in two ways. We show that (i) verbal instruction unaccompanied by a demonstration influences older children's exploration and (ii) that the impact of instruction on exploration depends on the exact nature of the instruction that children receive. When instruction does not conflict with children's intuitions about what they observe, it leads them to narrow their exploration to a subset of the various possibilities that they would have investigated on their own, allowing them to focus and restrict their exploration in an efficient fashion. By contrast, when older children are presented with information that conflicts with their intuitions, such information helps them to learn by prompting them to consider and test possibilities they would not have considered otherwise. It broadens the scope of their exploration. Indeed, older children in the counterintuitive testimony condition picked up more dolls (specifically the biggest and the smallest doll) than older children in the confirming testimony condition. To summarize, it appears that whether children are taught novel or counter-intuitive information, instruction influences exploration by focusing children's attention and exploration to the target of that instruction.

Third, our findings lend support to the claim that children's learning from testimony is a process of rational inference such that children are using their knowledge of the physical and social world to learn from testimony (Sobel & Kushnir, 2013). In our experiment, we demonstrated that children's reception of social information along with their beliefs about the relationship between size and weight shaped their decision to search for information and in turn, the empirical data that children gathered influenced their conclusions about the dolls. An interesting question is whether children's search for information also influenced their beliefs about the informant's future reliability. We might predict that, having discovered that the informant provided incorrect information, children who explored the doll in the counter-intuitive testimony condition would be less likely to endorse future testimony by that informant. Although, we did not gather data to answer this question, a recent experiment by Bridgers and colleagues (2016) revealed that when 4-year-old children observed physical evidence that contradicted an informant's testimony they were subsequently less trusting of that informant (see also Ronfard & Lane, 2017). Their findings suggest that in our experiment, the realization by some children that they had received incorrect testimony could have decreased their trust in the experimenter had they received additional testimony from her.

Older children's targeted exploration following counter-intuitive testimony opens up two questions for future research. First, are there conditions under which younger children would seek evidence following a counter-intuitive claim? For example, are younger children more likely to seek evidence when the task is simpler, when they are faced with a clearly unreliable informant, or when they are faced with a strongly counter-intuitive claim? Our experimental task may have been particularly demanding for preschool children because it required them to make a comparison across two dimensions (i.e., weight as well as height). In future studies, it will be important to find out if preschool children also fail to check an unexpected claim when a single dimension, such as quantity, is involved. Moreover, if younger children fail to investigate counterintuitive claims because of their greater trust in an experimenter's testimony relative to older children, then children's empirical testing of an experimenter's claim should increase when they are faced with a less reliable experimenter or when they are faced with a claim that more strongly clashes with their intuitions.

Second, what role does culture play in shaping children's search for evidence following counter-intuitive claims and in their willingness to report that evidence once it has been gathered? Chinese children are generally socialized towards minimizing social conflict rather than towards expressing their own beliefs and opinions (Chen & French, 2008; Markus & Kitayama, 1991). Consistent with these socialization goals, Chinese American and Chinese children are more likely to endorse counter-perceptual testimony provided by a consensus of multiple informants than are European American children – at least when asked to make these judgments publicly (Corriveau et al., 2013; Corriveau & Harris, 2010). However, they do not differ from their European American counterparts when making these judgments privately (Corriveau & Harris, 2010). Because we provided children in our experiment with the opportunity to explore the dolls in private, we suspect that the age-related differences in exploration that we observed will also be found in other cultures. However, we anticipate that children in the United States might be more willing to spontaneously report evidence conflicting with an adult's claim – a rare occurrence in our study, even among the elementary school children. Recall that overall very few children in the present study commented on the weight of the dolls after the return of E1. Thus, we anticipate that future work may show that cultural differences in the way that children in the U.S. and China are socialized lead to differences in how children interact with an adult but not in their search for evidence following a counter-intuitive claim. By implication, culture does not shape whether children seek evidence following a counterintuitive claim but rather whether and in what manner they disclose the information they have gathered to an adult, especially an adult who has made that claim.

In summary, young children are receptive to claims that defy their perceptions and intuitions. These claims allow them to quickly acquire many beliefs and practices that they would not be able to learn on their own. Nevertheless, when opportunities are available, elementary school children seize opportunities to evaluate counter-intuitive claims through empirical investigation and revise their judgments accordingly. This empirical stance implies that children's endorsement of counter-intuitive claims does not imply an unreflective acceptance of such claims and shows that, when learning about counter-perceptual and counterintuitive aspect of the physical world, instruction can lead children to explore hypotheses they would not have considered on their own.

References

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infants' ability to draw inferences about nonobvious object properties: Evidence from exploratory play. *Child Development, 64*, 711-728. doi: 10.1111/j.1467-8624.1993.tb02938.x

Bernard, S., Harris, P. L., Terrier, N., & Clément, F. (2015). Children weigh the number of informants and perceptual uncertainty when identifying objects. *Journal of Experimental Child Psychology, 136*, 70‑81. doi: 10.1016/j.jecp.2015.03.009Bonawitz, 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. doi:10.1016/j.cognition.2010.10.001

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

Bridgers, S., Buchsbaum, D., Seiver, E., Griffiths, T. L., & Gopnik, A. (2015). Children's causal inferences from conflicting testimony and observations. *Developmental Psychology, 52*, 9-18. http://dx.doi.org/10.1037/a0039830

Ceci, S. J., & Bruck, M. (2006). Children’s suggestibility: Characteristics and mechanisms. In R. V. Kail (Ed.), *Advances in Child Development and Behavior* (pp. 247–281). San Diego, CA: Elsevier. doi:10.1016/S0065- 2407(06)80009-1

Chan, C. C., & Tardif, T. (2013). Knowing better: the role of prior knowledge and culture in trust in testimony. *Developmental Psychology*, *49*, 591-601. doi: 10.1037/a0031336

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

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*, 148-153. doi: 10.1016/j.tics.2009.01.005

Chen, X., & French, D. C. (2008). Children's social competence in cultural context. *Annual Review of Psychology*, *59*, 591–616. doi: 10.1146/annurev.psych.59.103006.093606

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*, 1098-1120. doi: 10.1111/1467-8624.00081

Corriveau, K. H., & Harris, P. L. (2010). Preschoolers (sometimes) defer to the majority when making simple perceptual judgments. *Developmental Psychology*, *26*, 437-445. doi: 0.1037/a0017553

Corriveau, K. H., Kim, E., Song, G. & Harris, P. L. (2013). Young children's deference to a majority varies by culture. *Journal of Cognition and Culture*, *13*, 367-381. doi: 10.1163/15685373-12342099

Filippova, E., & Astington, J. W. (2008). Further development in social reasoning revealed in discourse irony understanding. *Child Development*, 79, 126–138. doi:10.1111/j.1467-8624.2007.01115.x

Fitneva, S. A., Lam, N. H. L., & Dunfield, K. A. (2013). The development of children's information gathering: To look or to ask? *Developmental Psychology*, *49*, 533–542. doi:10.1037/a0031326

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23, 183-209. doi:10.1016/0010-0277(86)90034-X

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*, 1532-1541. doi: 10.2307/1130693

Harris, P.L. (2012). *Trusting what you're told: How children learn from others*. Cambridge, MA: Belknap Press/Harvard University Press.

Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development*, *77*, 505–524. doi: 10.1111/j.1467-8624.2006.00886.x

Harris, P.L., Koenig, M.A., Corriveau, K.H. & Jaswal, V.K. (in press). Cognitive foundations of learning from testimony. *Annual Review of Psychology*.

Henrich, J., Heine, S. & Norenzayan, A. (2010). The weirdest people in the world? Behavioral and Brain Sciences, *33*, 61-83. doi:10.1017/S0140525X0999152X

Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, *21*, 1541–1547. doi:10.1177/095679761 0383438

Jaswal, V. K., Lima, O. K., & Small, J. E. (2009). Compliance, conversion, and category induction. *Journal of Experimental Child Psychology*, *102*, 182-195. doi: 10.1016/j.jecp.2008.04.006

Jaswal, V. K. (2010). Believing what you're told: Young children's trust in unexpected testimony about the physical world. *Cognitive Psychology*, *61*, 248-272. doi: 10.1016/j.cogpsych.2010.06.002

Jaswal, V. K., & Malone, L. S. (2007). Turning believers into skeptics: 3-year-olds' sensitivity to cues to speaker credibility. *Journal of Cognition and Development*, *8,* 263-283.

Jaswal, V. K., Perez-Edgar, K., Kondrad, R. L., Palmquist, C. M., Cole, C. A., & Cole, C. E. (2014). Can't stop believing: Inhibitory control and resistance to misleading testimony. *Developmental Science*, *17*, 1–12. doi:10.1111/desc.12187

Krogh-Jespersen, S., & Echols, C. H. (2012). The influence of speaker reliability on first versus second label learning. *Child Development*, *83*, 581–590. doi:10.1111/j.1467-8624.2011.01713.x

Koenig, M. A., & Echols, C. H. (2003). Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition*, *87,* 179-208.

Lane, J. D., & Harris, P. L. (2014). Confronting, representing, and believing counterintuitive concepts: Navigating the natural and the supernatural. *Perspectives on Psychological Science, 9,* 144-160. doi: 10.1177/1745691613518078

Lane, J. D., Harris, P. L., Gelman, S. A., & Wellman, H. M. (2014). More than meets the eye: Young children's trust in claims that defy their perceptions. *Developmental Psychology, 50*, 865-71. doi: 10.1037/a0034291

Lucas, A. J., Burdett, E. R., Burgess, V., Wood, L. A., McGuigan, N., Harris, P. L., & Whiten, A. (2016). The development of selective copying: children's learning from an expert versus their mother. *Child Development*. Advance online publication. doi: 10.1111/cdev.12711

Ma, L., & Ganea, P. A. (2010). Dealing with conflicting information: Young children's reliance on what they see versus what they are told. *Developmental Science, 13,* 151-160. doi: 10.1111/j.1467-7687.2009.00878.x

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224-253. doi: 10.1037/0033-295X.98.2.224

Mounoud, P., & Bower, T.G.R. (1974). Conservation of weight in infants. *Cognition, 3*, 29-40.

Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science*, 16, 385-390. doi: 10.1111/j.0956-7976.2005.01545.x

Peterson, C. C., Wellman, H. M., & Slaughter, V. (2012). The mind behind the message: Advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or Asperger syndrome. *Child Development*, *83*, 469–485. doi:10.1111/j.1467-8624.2011.01728.x

Piekny, J., Grube, D., & Maehler, C. (2014). The development of experimentation and evidence evaluation skills at preschool age. *International Journal of Science Education*, *36*, 334-354.

Poole, D. A., & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied, 7,* 27–50. doi:10.1037/1076-898X.7.1.27

Ronfard, S., & Lane, J. D. (2017). Preschoolers continually adjust their epistemic trust based on an informant's ongoing accuracy. *Child Development*. Advance online publication. doi: 10.1111/cdev.12720

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

Shafto, P., Goodman, N.D. & Frank, M.C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*, 341-351. doi: 10.1177/1745691612448481

Shneidman, L., Gweon, H., Schulz, L. E., & Woodward, A.L. (2016). Learning from others and spontaneous exploration: A cross-cultural investigation. *Child Development*, *87*, 723-735. doi: 10.1111/cdev.12502

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review*, *120*, 779 –797. http://dx.doi.org/10.1037/a0034191

Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concepts of size, weight, and density. *Cognition*, *21*, 177-237. doi:10.1016/0010-0277(85)90025-3

Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, *62*, 753-766.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*, 91-94. doi: 10.1126/science.aaa3799

Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, *79*, 866-881. doi: 10.1111/j.1467-8624.2008.01164.x

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

Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony young children trust inaccurate informants. *Developmental Science*, *17*, 443–451. doi:10.1111/desc.12134

Wimmer, H., Gruber, S., & Perner, J. (1984). Young children's conception of lying: Lexical realism—moral subjectivism. *Journal of Experimental Child Psychology*, *37*, 1-30. doi: 10.1016/0022-0965(84)90055-9