Developmental Psychology
2007, Vol. 43, No. 5, 1216–1226

0012-1649/07/$12.00 DOI: 10.1037/0012-1649.43.5.1216

Preschoolers Monitor the Relative Accuracy of Informants

Elisabeth S. Pasquini, Kathleen H. Corriveau, Melissa Koenig, and Paul L. Harris
Harvard University

In 2 studies, the sensitivity of 3- and 4-year-olds to the previous accuracy of informants was assessed. Children viewed films in which 2 informants labeled familiar objects with differential accuracy (across the 2 experiments, children were exposed to the following rates of accuracy by the more and less accurate informants, respectively: 100% vs. 0%, 100% vs. 25%, 75% vs. 0%, and 75% vs. 25%). Next, children watched films in which the same 2 informants provided conflicting novel labels for unfamiliar objects. Children were asked to indicate which of the 2 labels was associated with each object. Three-year-olds trusted the more accurate informant only in conditions in which 1 of the 2 informants had been 100% accurate, whereas 4-year-olds trusted the more accurate informant in all conditions tested. These results suggest that 3-year-olds mistrust informants who make a single error, whereas 4-year-olds track the relative frequency of errors when deciding whom to trust.

Keywords: preschoolers, accuracy, informants, testimony

Much of the information children are exposed to cannot be learned through direct experience. To learn about the shape of the Earth, the existence of germs, or the role of the brain in mental functioning, children must rely on the testimony of others (Harris & Koenig, 2006, in press). However, it is unlikely that children accept all they are told, particularly when they receive conflicting testimony from different speakers. Children are likely to employ heuristics to assess whose input is more reliable. Three recent studies show that the past accuracy of informants is one of the cues that preschool children use when deciding whose testimony to accept (Clément, Koenig, & Harris, 2004; Jaswal & Neely, 2006; Koenig, Clément, & Harris, 2004). When presented with one informant who consistently made errors in naming familiar objects and another informant who was consistently correct, preschoolers were more likely to accept information from the accurate as opposed to the inaccurate informant when the informants subsequently provided conflicting information.

In two recent studies, the performance of 3- and 4-year-olds has been compared. Koenig and Harris (2005) found that 4-year-olds, but not 3-year-olds, were above chance in choosing to accept information from an informant who had consistently been accurate as opposed to an informant who had consistently been inaccurate. Similarly, when Birch, Luca, Frampton, Vauthier, and Bloom (2005) presented children with one informant who consistently provided accurate information about the function of objects and a second informant who consistently provided inaccurate information, 4-year-olds, but not 3-year-olds, used this information when deciding which informant to rely upon for information about object labels. However, when Koenig and Harris presented 3- and 4-year-olds with one consistently accurate informant and one consistently ignorant informant (who acknowledged not knowing the names of familiar objects), both 3- and 4-year-olds were above chance in seeking and endorsing information from the accurate rather than the ignorant informant.

These results raise two issues that are explored in this article. First, when children display selective trust, what strategy do they adopt in monitoring informants? Second, why are 3-year-olds less successful in monitoring for accuracy than 4-year-olds? We consider each of these issues in turn. Previous studies set up a contrast between a consistently accurate informant and a consistently inaccurate (or ignorant) informant, yet such consistent differences between informants will rarely occur in children's everyday experience. Instead, informants generally provide a mix of accurate and inaccurate information with relatively stable individual differences in their overall degree of accuracy. There are at least three strategies that children could employ in determining which informants to trust under such conditions. First, children could monitor for accuracy, trusting informants who have made an accurate claim on at least one occasion. Second, children could monitor for inaccuracy, trusting informants who have not made an inaccurate claim. Finally, children could engage in a form of statistical monitoring, rather than simply treating informants as accurate or inaccurate.

A statistical monitoring strategy requires that children track the relative frequency of event types. For example, children might track the number of times that each informant has made an error and place more trust in the informant who has made fewer errors. Although this strategy might appear to be challenging, a number of studies have shown that young children are able to use statistical operations in learning. For example, 8-month-old infants can extract statistical patterns from visual arrays (Kirkham, Slemmer, & Johnson, 2002). Eight-month-old infants can also parse the speech stream into segments based on the statistical relationships between neighboring speech sounds (Aslin, Saffran, & Newport, 1998;

Elisabeth S. Pasquini, Kathleen H. Corriveau, Melissa Koenig, and Paul L. Harris, Harvard Graduate School of Education, Harvard University.

Melissa Koenig is now at the Institute of Child Development, University of Minnesota.

This article is based on work supported under a National Science Foundation Graduate Research Fellowship to Elisabeth S. Pasquini. We are grateful to all of the children who participated in this study.

Correspondence concerning this article should be addressed to Paul L. Harris, 503A Larsen Hall, Harvard Graduate School of Education, Appian Way, Cambridge, MA 02138. E-mail: paul_harris@gse.harvard.edu

Saffran, Aslin, & Newport, 1996). Finally, preschool children have been found to apply statistical calculations in their causal learning (Kushnir & Gopnik, 2005). Taken together, these findings show that young children are capable of making statistical calculations in multiple domains. Accordingly, they may use such a strategy in deciding whom to trust.

Turning to the age change between 3 and 4 years, Koenig and Harris (2005) suggested that 3-year-olds' inferior performance when choosing between accurate versus inaccurate speakers, but not between accurate versus ignorant speakers, might be explained in terms of their limited understanding of false belief. Three-year-olds typically lack the ability to represent the false beliefs that underlie mistaken utterances and actions (Wellman, Cross, & Watson, 2001). However, 3-year-olds are generally more accurate in attributing ignorance than in attributing false beliefs (Hogrefe, Wimmer, & Perner, 1986; Perner & Wimmer, 1988). Thus, 3-year-olds may not know how to interpret an informant who consistently misnames objects but do understand an informant who consistently acknowledges ignorance. An alternative explanation for 3-year-olds' inferior performance is that despite some ability to interpret misnaming, 3-year-olds found it difficult to differentiate and keep track of the accurate versus the inaccurate informant. On this hypothesis, 3-year-olds might show selective trust if the difference between the two informants were more obvious or salient. Experiment 1 was designed to test each of these two hypotheses.

In summary, Experiment 1 had three main goals. One goal was to check which strategy children use to determine which of two informants to trust. Children were tested in three different conditions. The 100% versus 0% condition replicated Experiment 1 in Koenig and Harris (2005), with minor modifications. Across four trials, one informant was consistently accurate when labeling common objects (100% correct), and one informant was consistently inaccurate (0% correct). In the 75% versus 0% condition, one informant was mostly accurate (75% correct), and one informant was consistently inaccurate (0% correct). Finally, in the 100% versus 25% condition, one informant was consistently accurate (100% correct), and the other was mostly inaccurate (25% correct).

Regardless of their strategy, we anticipated that 4-year-olds would display selective trust in the 100% versus 0% condition. However, if 4-year-olds were simply monitoring for any display of accuracy, they should display selective trust in the 75% correct versus 0% correct condition but not in the 100% correct versus 25% correct condition, as both informants labeled at least one object correctly in the latter condition. On the other hand, if 4-year-olds were simply monitoring for any display of inaccuracy, they should display selective trust in the 100% versus 25% condition but not in the 75% versus 0% condition, as both informants labeled at least one object incorrectly in the latter condition. Finally, if 4-year-olds were accurate across all three conditions, it would suggest that they are able to use the more challenging, statistical strategy.

A second goal of Experiment 1 was to examine the relationship between selective trust and false belief understanding. As noted earlier, Koenig and Harris (2005) suggested that 3-year-olds might experience difficulty in interpreting false labels because they do not yet understand the false beliefs that may motivate them (Gopnik & Astington, 1988; Wellman et al., 2001). To determine whether selective trust is associated with an understanding of false beliefs, we included a standard assessment of false belief understanding. If performance on this task is predictive of selective trust after controlling for age, this would support the proposal that selective trust depends on children's ability to attribute mistaken utterances to false beliefs.

The third goal of Experiment 1 was to check whether 3-year-olds' performance would improve if the two informants were easier to differentiate and keep track of. Accordingly, we modified the paradigm used by Koenig and Harris (2005) in three ways. First, the experimenter referred to each informant by the color shirt she was wearing and used this name during each trial. Second, we did not vary the position of the informants between trials. Finally, we increased the amount of exposure to each informant by increasing the number of familiar objects that each informant labeled (from three to four).

Experiment 1

Method

Participants. Forty-one children participated in this study: 21 three-year-olds (M = 3 years 6 months; range = 3 years 0 months to 3 years 11 months; 10 girls) and 20 four-year-olds (M = 4 years 5 months; range = 4 years 1 month to 5 years 1 month; 7 girls). All children were recruited from a child-care center in Buffalo, New York, serving a broad socioeconomic range. Approximately 50% of children were White, 27% were Eastern Asian, and the remaining children were primarily Southeast Asian and African American.

False belief task. At the start of testing, each child participated in an unexpected contents false belief task. Children were first presented with a crayon box and asked what they thought was inside the box. All children responded "crayons" or "markers." Children were then invited to open the box. The box contained a set of birthday candles. With the help of the experimenter, the candles were returned to the box, and the box was closed. Next, children were asked three test questions:

(1) "When you first saw the box, before we opened it, what did you think was inside? Did you think there were crayons inside it, or did you think there were candles inside it?"

(2) "If one of your friends came in here, who had never seen this box before, what would she think was inside? Would she think there were candles inside or would she think there were crayons inside?"

(3) "Here's a bunny [experimenter presents puppet]. Let's pretend this bunny has been asleep all the time we've been talking. Say we showed him this box. Would Bunny think there were crayons inside or would he think there were candles inside?"

Additionally, children were asked one control question immediately following the first test question: "What's really inside the box? Are there candles inside the box or are there crayons inside the box?" All children answered this question correctly. Scores for this task were out of a possible total of three correct.

Selective trust tasks. Children then proceeded to the selective trust tasks. Each child was tested in all three conditions, with the order of conditions systematically varied across participants.

Materials. For each of the three conditions, we created a short film; each film comprised eight clips. The same three actors were shown in all eight clips of a given film (although each film featured a different set of three actors). In each film, two female

actors wearing different solid-colored shirts (Film 1: pink and yellow; Film 2: white and purple; Film 3: green and red) were seated at a table. In each film, the two female actors were similar in age and appearance. Each clip began with a male actor standing behind two female actors and placing an object on the table between them. On accuracy trials, the four objects were familiar (e.g., ball, shoe; see Table 1 for a full list of objects). On test trials, the four objects were novel (e.g., a toilet flapper, a sprinkler head; see Table 2 for a full list of test objects). Before each test clip was played, children were presented with a still photograph of the relevant object and were asked if they knew what the object was called. The order of trials within accuracy and test periods was maintained across participants, as shown in Tables 1 and 2.

Design and procedure. To introduce the task, the experimenter pointed to a still frame from the film and said,

> I've got these two friends. See? One has a pink [white, green] shirt and one has a yellow [purple, red] shirt. They're going to show you some things and tell you what they are called. I want you to listen very carefully and then I'm going to ask you some questions. Let's watch.

Accuracy trials. On each accuracy trial, children were presented with a picture of a familiar object and then watched a video clip of three actors and the familiar object. Trials began when the male actor placed the object on the table between the two female actors and asked one actor, "Can you tell me what this is called?" The first female actor responded by saying, "That's a _____," and the same question was posed to the second female actor, who responded with a different label, saying, "That's a _____." In each film, the order in which the actors were asked to name the familiar object alternated across the four video clips. In every clip, object names were matched for age of acquisition (Fenson et al., 1994). The mean age of acquisition of object names did not differ across films. Table 1 lists all object names used in the accuracy trials.

In Film 1 (100% vs. 0% condition), one actor named all four objects correctly (100% correct). For example, when presented with a brush, the accurate informant said, "That's a brush." The other actor named all four objects incorrectly (0% correct). For example, when presented with a brush, the inaccurate informant said, "That's a plate." After both of the actors named the object, the experimenter paused the video and said, "The girl with the pink shirt said it's a ____ and the girl with the yellow shirt said it's a ____. What do you think it's called?" For half of the participants, the actor with the yellow shirt was 100% correct. For the other half, the actor with the pink shirt was 100% correct.

In Film 2 (75% vs. 0% condition), one actor named three of the four objects correctly and one object incorrectly (75% correct). The other actor named all four objects incorrectly (0% correct). On the trial in which both actors named the familiar object incorrectly, the actors assigned different incorrect names to the object. Following each video clip, the experimenter asked the same question as in Film 1. For half of the participants, the purple-shirted actor was 75% correct, and for the other half, the white-shirted actor was 75% correct. The particular trial in which both actors incorrectly named the object varied randomly across participants.

In Film 3 (100% vs. 25% condition), one actor named all four objects correctly (100% correct). The other actor named one object correctly and three objects incorrectly (25% correct). Following each video clip, the experimenter asked the same question as in Films 1 and 2. For half of the participants, the green-shirted actor was 100% correct, and for the other half, the red-shirted actor was 100% correct. The particular trial in which both actors correctly named the object varied randomly across participants.

Test trials. Following the fourth accuracy trial, all participants were shown four successive still photographs of novel objects and four corresponding video clips. These clips began with the male actor placing a novel object on the table between the two female actors and then asking one actor, "Can you tell me what this is called?" The first female actor responded by producing a novel name (e.g., "That's a *toma*") and the same question was posed to the second female actor, who produced a different novel name (e.g., "That's a *danu*"). In each film, the order in which the actors were asked to name the novel object alternated across the four video clips. A full list of the novel names and descriptions of the novel objects can be found in Table 2. During this test period, children were asked four types of test questions.

In the two explicit judgment trials, the experimenter referred to a still frame of the video and asked, "One of these people was not very good at answering these questions. Which person was not

Table 1
Stimuli Used in Accuracy Trials for Experiment 1

Condition	Familiar object	Accurate label	Inaccurate label
100% versus 0%	Spoon	"That's a spoon"	"That's a duck"
	Bottle	"That's a bottle"	"That's an apple"
	Brush	"That's a brush"	"That's a plate"
	Doll	"That's a doll"	"That's a cup"
75% versus 0%	Bowl	"That's a bowl"	"That's a lion"
			"That's a clock"
	Shovel	"That's a shovel"	"That's a penny"
			"That's a towel"
	Ball	"That's a ball"	"That's a cookie"
			"That's a dog"
	Shoe	"That's a shoe"	"That's a truck"
			"That's a nose"
100% versus 25%	Bear	"That's a bear"	"That's a hat"
	Boat	"That's a boat"	"That's a tree"
	Book	"That's a book"	"That's a key"
	Phone	"That's a phone"	"That's a fork"

Table 2
Stimuli Used in Test Trials of Experiment 1

Condition	Novel object	Informant 1 name	Informant 2 name
100% versus 0%	Gray rubber squeegie	"That's a toma"	"That's a gobi"
	Metallic square deer warner	"That's a modi"	"That's a danu"
	Yellow plastic sprinkler attachment	"That's a dax"	"That's a wug"
	Blue and white plastic toilet flapper	"That's a fep"	"That's a riff"
75% versus 0%	Metallic cocktail strainer	"That's a plick"	"That's a lorg"
	Metallic and black cocktail pourer	"That's a merval"	"That's a feppin"
	Black and gray knee pad	"That's a cham"	"That's a roke"
	Silver bathroom door hook	"That's a nevi"	"That's a mogo"
100% versus 25%	Black toilet bulb	"That's a neem"	"That's a zav"
	Metallic lemon juicer	"That's a niddy"	"That's a gabber"
	Wooden lemon juicer	"That's a norp"	"That's a liff"
	Gold and red sprinkler head	"That's a terval"	"That's a blicket"

very good at answering these questions?" This explicit judgment question was asked after the fourth accuracy trial and again at the end of the experiment, directly after the fourth test trial.

The four ask trials occurred prior to the viewing of each video clip. Children were presented with a still photograph of a novel object and were asked, "Do you know what this is called?" Children were given a chance to reply and were then asked, "I bet one of these people can help us find out what it is called. Which person would you like to ask, the person with the pink [white, green] shirt or the person with the yellow [purple, red] shirt?" Children who claimed to know the name of the novel object were told, "Actually, I don't think that's what it is called. I bet one of these people can help us find out what it is called. Which person would you like to ask, the person with the pink [white, green] shirt or the person with the yellow [purple, red] shirt?"

The four endorsement trials occurred after the children viewed each video clip. The experimenter paused the video and questioned children in a similar manner to that used in the accuracy trials (e.g., "The girl in the pink shirt said it's a *danu* and the girl in the yellow shirt said it's a *modi*. What do you think it's called, a *danu* or a *modi*?").

Results

We first report on children's performance during the accuracy trials. We then assess children's performance on the three test questions in each condition against chance. Next, we analyze children's performance on the critical trust probes, namely the ask and endorse questions. We then describe children's performance on the false belief task. Finally, we examine the relationship between performance on the false belief task and overall performance on the test questions.

Performance during the accuracy trials. Every 3- and 4-year-old accurately chose the correct name for the familiar objects in the 100% versus 0% and the 100% versus 25% conditions.

However, six 3-year-olds and two 4-year-olds erred on one accuracy trial in the 75% versus 0% condition. One 3-year-old endorsed the incorrect name on an accuracy trial with one accurate and one inaccurate informant, agreeing with the inaccurate informant that the object (a bowl) was a clock. This child was removed from the sample. The remaining seven children's errors occurred on the single trial in which both informants incorrectly named the familiar object. Children generally chose to endorse the informant who had previously been accurate.

Comparisons with chance for three test questions in the three conditions. Proportion correct and comparisons to chance for the explicit judgment, ask, and endorse questions are shown in Table 3 as a function of age and condition. Three-year-olds performed above chance on all three question types in the 100% versus 0% condition and the 100% versus 25% condition. However, they performed at chance on all three question types in the 75% versus 0% condition. In contrast, 4-year-olds performed above chance on all three question types across all three conditions.

Performance on ask and endorse questions. In order to assess performance on the ask and endorse questions, we calculated a three-way analysis of covariance (ANCOVA), with age group (3-, 4-year-olds) as the between-subjects variable and question type (ask, endorse) and condition (100% vs. 0% correct, 75% vs. 0% correct, 100% vs. 25% correct) as the within-subjects variables; scores on the explicit judgment questions were entered as a covariate. This ANCOVA revealed main effects of age group, $F(1, 35) = 14.77$, $p < .001$, $\eta^2 = .14$, and of explicit judgment scores, $F(1, 36) = 4.93$, $p < .05$, $\eta^2 = .12$, but no main effects of condition, $F(2, 72) = 0.91$, *ns*, or of question type, $F(1, 36) = 2.31$, *ns*. No other significant main effects or interactions were detected. Thus, performance on ask and endorse questions varied as a function of performance on the initial explicit judgment questions. However, even with performance on the explicit judgment questions included as a covariate, 4-year-olds still performed better on the ask and endorse questions than did 3-year-olds.

Performance on false belief task. Scores for proportion correct (maximum = 3) on the false belief test questions are displayed in Table 3. Three-year-olds performed below chance on the false belief questions, indicating that the majority of these children believed that all uninformed observers would know that the crayon box contained candles. Four-year-olds' performance was not significantly different from chance performance. Two 3-year-olds and four 4-year-olds received a score of 3, three 3-year-olds and eight 4-year-olds received a score of 2, and 15 3-year-olds and eight 4-year-olds received a 1 or 0. In summary, the majority of 3-year-olds did not attribute false beliefs to either themselves or others, whereas 4-year-olds exhibited a more mixed pattern of performance.

Table 3
Proportion Correct and Comparisons With Chance for Experimental Measures in Experiment 1

	3-year-olds			4-year-olds		
Question	Proportion	*t*(19)	95% CI	Proportion	*t*(19)	95% CI
			100% versus 0% correct			
Explicit judgment	.71 (.08)	3.25*	.55, .87	.90 (.05)	7.55***	.80, .99
Ask	.70 (.06)	3.35**	.58, .82	.84 (.04)	8.10***	.76, .92
Endorse	.67 (.06)	2.80*	.55, .79	.88 (.04)	8.82***	.80, .96
			75% versus 0% correct			
Explicit judgment	.48 (.09)	0.00	.30, .65	.78 (.08)	3.24**	.62, .94
Ask	.61 (.07)	1.23	.47, .75	.85 (.06)	6.29***	.73, .97
Endorse	.50 (.07)	−0.37	.36, .64	.80 (.05)	5.64***	.70, .90
			100% versus 25% correct			
Explicit judgment	.85 (.05)	5.94***	.75, .95	.92 (.04)	9.80***	.84, .99
Ask	.65 (.07)	2.11*	.52, .79	.79 (.07)	4.20***	.65, .93
Endorse	.68 (.05)	3.47**	.58, .78	.80 (.07)	4.33***	.66, .94
			False belief task			
False belief total	.21 (.08)	3.49***	.05, .37	.53 (.08)	0.43	.37, .69

Note. Standard errors are in parentheses. CI = confidence interval.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Relationship between false belief and test questions. To examine the extent to which variation in performance on test questions could be accounted for by false belief score, controlling for the effects of age, we calculated an ANCOVA, with false belief score (out of 3) as the between-subjects variable and question type (explicit judgment, ask, endorse) and condition (100% vs. 0% correct, 75% vs. 0% correct, 100% vs. 25% correct) as the within-subjects variables, and with age (in months) as a covariate. The ANCOVA results indicated a significant effect of age, $F(1, 37) = 7.00$, $p < .05$, $\eta^2 = .175$, but no significant effect of false belief score, $F(3, 35) = 1.02$, *ns*. No other significant main effects or interactions were detected.

To further examine whether tracking the accuracy of informants requires an understanding of false beliefs, we examined the performance of the 18 children who performed poorly on the false belief task (defined as children who answered all false belief questions incorrectly). Despite these children's poor understanding of false belief, the sum of their scores on explicit judgment, ask, and endorse questions across all three conditions (maximum = 30) was significantly above chance ($M = 19.67$, $SD = 4.96$), $t(17) = 3.99$, $p < .001$, as was their total score (maximum = 10) on the 100% versus 0% condition ($M = 7.00$, $SD = 1.60$), $t(17) = 5.27$, $p < .001$, and the 100% versus 25% condition ($M = 6.94$, $SD = 2.21$), $t(17) = 3.74$, $p < .01$. These children's performance did not differ from chance in the 75% versus 0% condition ($M = 5.78$, $SD = 2.71$), $t(17) = 1.22$, *ns*. In summary, we found little evidence of a relationship between false belief understanding and success on test questions.

Discussion

Overall, there were three main findings of interest. First, both 3- and 4-year-olds generally performed above chance in identifying the less accurate informant and in using this information to decide which informant to ask and endorse. However, 3-year-olds were at chance across all three questions in the 75% versus 0% condition. Second, 4-year-olds performed better than 3-year-olds on ask and endorse questions, even when performance on the explicit judgment questions was included as a covariate. Finally, children's understanding of false beliefs did not predict their performance on test trials, after controlling for age.

One of our goals was to assess whether children who display selective trust rely on an accuracy strategy (trusting any informant who has provided at least some accurate information in the past), an inaccuracy strategy (trusting only informants who have never provided inaccurate information), or a statistical strategy (trusting the more accurate informant). We consider the likely strategy of each age group in turn.

Three-year-olds' above-chance performance in the 100% versus 0% condition and the 100% versus 25% condition, together with their chance performance in the 75% versus 0% condition, suggests that they use the inaccuracy strategy. Thus, in the 100% versus 0% condition and the 100% versus 25% condition, they trusted the informant who made no errors but mistrusted the informant who had made an error. By contrast, in the 75% versus 0% condition, they did not display selective trust because each informant had made an error. This interpretation is also consistent with our observation that some children answered the explicit judgment question in the 75% versus 0% condition by saying that both speakers had been "not very good" at answering the questions. By implication, 3-year-olds were relatively unforgiving of the single error made by the more accurate informant in the 75% versus 0% condition. To the extent that she had made at least one error, she too was judged not very good at answering the questions and hence deemed untrustworthy.

Four-year-olds' above-chance performance in all conditions suggests that they were adopting neither the accuracy nor the inaccuracy strategy. They did not simply focus on whether a given informant had ever made an accurate or an inaccurate claim. Thus, although both informants made inaccurate claims in the 75% versus 0% condition, 4-year-olds trusted the more accurate of the two. Similarly, although both informants made accurate claims in the 100% versus 25% condition, 4-year-olds trusted the more accurate of the two. These findings suggest that 4-year-olds were using a form of statistical monitoring. They tracked the frequency of accurate versus inaccurate claims made by each informant and invested greater trust in the informant who had been accurate more often. However, it is important to note that 4-year-olds' success in all three conditions could also be explained by a two-rule system. Thus, if 4-year-olds made decisions about which informant to trust using the rules (a) do not trust individuals who are always inaccurate and (b) do trust individuals who are always accurate, they would succeed in all three conditions without employing a statistical strategy. This issue was further examined in Experiment 2.

The second goal of Experiment 2 was to examine false belief understanding as a potential mechanism underlying the development of selective trust between the ages of 3 and 4. We found that 4-year-olds outperformed 3-year-olds on the false belief task, with 3-year-olds performing significantly below chance and 4-year-olds performing at chance levels. However, analyses controlling for age did not detect a relationship between false belief understanding and success in the test trials. Furthermore, children who answered every false belief question incorrectly were still above chance in demonstrating selective trust in the 100% versus 0% condition and the 100% versus 25% condition, suggesting that selective trust does not require an understanding of false belief.

The third goal of Experiment 1 was to examine whether 3-year-olds would display selective trust in a modified version of Experiment 1, reported by Koenig and Harris (2005). Our task differed from the experiment of Koenig and Harris in three ways, each aimed at making it easier for children to differentiate and keep track of the two informants. First, the experimenter consistently referred to each informant by the color shirt that she was wearing (e.g., "The girl with the pink shirt said it's a ____ and the girl with the yellow shirt said it's a ____. What do you think it's called?"). Second, the position of the informants remained constant between trials for individual children. Finally, we increased the number of accuracy trials (from three to four). In contrast to the 3-year-olds in Experiment 1 of Koenig and Harris, whose performance did not differ from chance levels, our 3-year-olds performed above chance levels in both the 100% versus 0% condition and the 100% versus 25% condition. Nevertheless, we also replicated the finding of Koenig and Harris that selective trust develops substantially between the ages of 3 and 4. Four-year-olds displayed greater selective trust than did 3-year-olds in all three conditions. Thus, the earlier conclusions of Koenig and Harris do stand but with important qualifications. Although, as they claimed, selective trust improves between the ages of 3 and 4, this is not because 3-year-olds are incapable of selective trust. So long as the inaccuracy strategy can be applied (i.e., trust only informants who have never provided inaccurate information) and so long as the two informants can be easily discriminated, 3-year-olds display selective trust at above-chance levels.

As noted earlier, some children, primarily in the 3-year-old group, initially balked at answering the explicit judgment question in the 75% versus 0% condition, commenting that both informants were not very good. On the basis of this observation, we hypothesized that 3-year-olds' poor performance in the 75% versus 0% condition may have been due, in part, to confusion resulting from the wording of the explicit judgment question. In order to test this hypothesis, we reworded the explicit judgment questions in Experiment 2.

In conclusion, the results suggest that 3- and 4-year-olds are alert to whether someone has made a false claim. They remember such errors, subsequently judge that the person is not good at answering questions, and when faced with conflicting claims, avoid seeking and accepting information from such an informant. The more systematic selectivity of 4-year-olds suggest that they use a statistical strategy (e.g., by tracking which informant has made the greater number of errors), whereas 3-year-olds use an inaccuracy strategy (by tracking which informant has made no errors). However, it is also possible that 4-year-olds' success was not based on a statistical strategy but on a two-rule strategy.

Experiment 2

Experiment 2 had two primary purposes. The first was to further explore the nature of the strategy underlying 4-year-olds' success in Experiment 1. Experiment 2 included a condition in which one informant was 75% accurate and the other was 25% accurate. If 4-year-olds monitor how often each informant has been inaccurate, they should determine that the 75% correct informant is accurate more often than the 25% correct informant. Hence, they should successfully demonstrate selective trust. Conversely, if 4-year-olds are employing a two-rule strategy, they should be indiscriminate in this condition. Because neither informant is 100% accurate or 100% inaccurate, neither of the two rules can be applied.

The second goal of Experiment 2 was to determine whether the poor performance of 3-year-olds in the 75% versus 0% condition of Experiment 1 was a result of the wording of the explicit judgment questions. Some children, particularly in the 3-year-old group, objected to the explicit judgment question, claiming that both informants were not very good. Performance on the explicit judgment questions might improve in a 75% versus 0% condition if they were reworded such that it was possible for children to characterize both informants as good or not good but also to differentiate between them. To test this hypothesis, we administered a 75% versus 0% condition, as well as a 75% versus 25% condition, but with reworded explicit judgment questions. Instead of simply asking which of the two informants was not very good (as we did in Experiment 1), we now asked a series of three questions. First, we asked children to evaluate one of the two informants as good or not very good at answering questions. Then, we asked children to make the same judgment about the second informant. Finally, we asked children which of the two informants was better at answering the questions.

Method

Participants. Fifty-seven children participated in this study: 25 three-year-olds (M = 3 years 7 months; range = 3 years 2 months to 3 years 11 months; 10 girls) and 32 four-year-olds (M = 4 years 6 months; range = 4 years 0 months to 5 years 0 months; 15 girls). All children were recruited from child-care centers in Cambridge, Massachusetts, and the surrounding area serving pre-

dominantly middle-class families. Approximately 83% of children were White, 10% were Eastern Asian, and the remaining children were primarily Southeast Asian and African American.

Materials. For each of the two conditions, we created a short film. In the 75% versus 0% condition, the film was the same film used in the 75% versus 0% condition of Experiment 1. In the 75% versus 25% condition, the film was similar to those used in Experiment 1, except that one informant was correct in 75% of accuracy trials and the other informant was correct in 25% of accuracy trials.

As in Experiment 1, children were presented with a still photograph of the relevant object before each clip was played. The order of trials within accuracy and test periods was maintained across participants, as shown in Tables 4 and 5, respectively.

Design and procedure. Each child was tested in both conditions, with the order of conditions systematically varied across participants. The instructions used to introduce the task were identical to those used in Experiment 1.

Accuracy trials. On each accuracy trial, children were first presented with a picture of a familiar object and then watched a video clip identical in format to those in Experiment 1, depicting three actors and the familiar object. Film 1 (75% vs. 25% condition) used the same clips found in Film 1 of Experiment 1, but these clips were edited such that one actor named three of the four objects correctly (75% correct) and the other actor named one of the four objects correctly (25% correct). The position of the clip in which the 75% correct informant made an error and the 25% correct informant was accurate was varied systematically across participants. In each clip, one actor named the object correctly, and the other actor named the object incorrectly (i.e., there were no clips in which both actors were correct or in which both actors made errors). The same procedure used in Experiment 1 was used in Experiment 2 for both films. Film 2 (75% vs. 0% condition) was identical to Film 2 of Experiment 1.

Test trials. The overall format of the test trials was similar to that of the test trials in Experiment 1. Thus, the wording of the ask and endorse questions was identical to the wording in Experiment 1. However, the explicit judgment questions were reworded in order to probe children's appraisal of the two informants more carefully. In the two explicit judgment trials, the experimenter referred to a still frame of the video and indicated one of the two informants, asking, "Was the girl with the _____ shirt good at answering the questions or was she not very good at answering the questions?" The same two-alternative forced-choice question was then asked of the other informant. The order of the forced-choice alternatives was varied across participants. Next, children were asked, "Who was better at answering the questions: the girl in the _____ shirt or the girl in the _____ shirt?" The order of these two alternatives was also counterbalanced across participants. These explicit judgment questions were asked after the final accuracy trial and again at the end of the experiment, after the final test trial.

Table 4
Stimuli Used in Accuracy Trials for Experiment 2

Condition	Familiar object	Accurate name	Inaccurate name
75% versus 25%	Spoon	"That's a spoon"	"That's a duck"
	Bottle	"That's a bottle"	"That's an apple"
	Brush	"That's a brush"	"That's a plate"
	Doll	"That's a doll"	"That's a cup"
75% versus 0%	Bowl	"That's a bowl"	"That's a lion"
			"That's a clock"
	Shovel	"That's a shovel"	"That's a penny"
			"That's a towel"
	Ball	"That's a ball"	"That's a cookie"
			"That's a dog"
	Shoe	"That's a shoe"	"That's a truck"
			"That's a nose"

Results

Performance during the accuracy trials. Most children accurately chose the correct name for the familiar objects in both conditions, with the exception of 3 three-year-olds and 6 four-year-olds. Of these 9 children, 1 made an error in the 75% versus 25% condition, 7 made errors in the 75% versus 0% condition, and 1 made errors in both conditions. Five of the children who made errors erred by endorsing the incorrect name on an accuracy trial with one accurate and one inaccurate informant (e.g., agreeing with the inaccurate informant that the object [a spoon] was a duck). These 5 children were removed from the sample. The remaining 4 children's errors occurred on the trial of the 75% versus 0% condition in which both informants incorrectly named the familiar object. All 4 children chose to endorse the informant who had previously been accurate.

Analysis of explicit judgment questions. On the basis of the results from Experiment 1, we hypothesized that children might have had difficulty with the explicit judgment question in the 75% versus 0% condition because they did not characterize one informant as good and the other as poor but instead characterized both informants as poor. To test this hypothesis, we included a new pair of explicit judgment questions, asking children to characterize each of the informants as good or not very good. Table 6 displays the pattern of children's characterizations on these two questions, along with the final explicit judgment question, by age and condition. Inspection of Table 6 shows that children offered more good judgments for the more accurate informant than for the less accurate informant.

To check this conclusion, we calculated a three-way analysis of variance (ANOVA), with age group (3-, 4-year-olds) as the between-subjects variable and informant accuracy (the more accurate informant, the less accurate informant) and condition (75% vs. 25% correct, 75% vs. 0% correct) as the within-subjects variables for good judgment scores. This ANOVA revealed a main effect of informant accuracy, $F(1, 56) = 37.98$, $p < .001$, $\eta^2 = .40$. No other significant main effects or interactions were found. On the basis of this result, we determined that, because children appropriately characterized one informant as good and the other as not very good, answers to the first two questions could be included in the total explicit judgment score along with scores on the final explicit judgment question.

Comparisons with chance. Next, we compared children's performance with chance for all three question types (i.e., explicit judgment, ask, endorse) in both conditions. The results of these comparisons can be found in Table 7. With the exception of their explicit judgment performance in the 75% versus 25% condition, 3-year-olds' performance did not differ significantly from chance in either of the two conditions. In contrast, 4-year-olds were able to explicitly identify the more accurate informant in both conditions. Four-year-olds were also selective in asking and endorsing the more accurate informant in the 75% versus 0% condition. In

Table 5
Stimuli Used in Test Trials of Experiment 2

Condition	Novel object	Informant 1 name	Informant 2 name
75% versus 25%	Gray rubber squeegie	"That's a toma"	"That's a gobi"
	Metallic square deer warner	"That's a modi"	"That's a danu"
	Yellow plastic sprinkler attachment	"That's a dax"	"That's a wug"
	Blue and white plastic toilet flapper	"That's a fep"	"That's a riff"
75% versus 0%	Metallic cocktail strainer	"That's a plick"	"That's a lorg"
	Metallic and black cocktail pourer	"That's a merval"	"That's a feppin"
	Black and gray knee pad	"That's a cham"	"That's a roke"
	Silver bathroom door hook	"That's a nevi"	"That's a mogo"

the 75% versus 25% condition, 4-year-olds were selective in endorsing claims made by the more accurate informant but not when asking for information.

Performance on ask and endorse questions. In order to assess performance on ask and endorse questions, we calculated a three-way ANCOVA, with age group (3-, 4-year-olds) as the between-subjects variable and question type (ask, endorse) and condition (75% vs. 25% correct and 75% vs. 0% correct) as the within-subjects variables; explicit judgment scores were included as a covariate. This ANCOVA revealed main effects of age group, $F(1, 49) = 4.24$, $p < .05$, $\eta^2 = .08$, and explicit judgment, $F(1, 49) = 11.44$, $p < .001$, $\eta^2 = .19$. No other significant main effects or interactions were found.

In summary, performance on ask and endorse questions varied as a function of performance on the initial explicit judgment questions. Nevertheless, even with performance on the explicit judgment questions included as a covariate, 4-year-olds still performed better on ask and endorse questions than did 3-year-olds.

Discussion

Experiment 2 replicated Experiment 1 in showing that 4-year-olds display selective trust when one informant is 75% accurate and the other is 0% accurate. Four-year-olds were also able to differentiate between an informant who was 75% accurate and an informant who was 25% accurate and preferred to seek information from the more accurate informant. By contrast, 3-year-olds did not show selective trust in either condition, despite modifications to the explicit judgment questions.

Four-year-olds could have succeeded in all three conditions of Experiment 1 either by monitoring the relative accuracy of the informants or by adhering to two simple rules: (a) do not trust individuals who are always inaccurate and (b) do trust individuals who are always accurate. To determine which strategy 4-year-olds use, we included a 75% accurate versus 25% accurate condition in Experiment 2. If 4-year-olds monitor the relative accuracy of informants, they would successfully demonstrate selective trust in the 75% versus 25% condition, but if 4-year-olds instead rely on a two-rule strategy, they would fail in this condition. In Experiment 2, we found that 4-year-olds responded selectively in the 75% versus 25% condition, with respect to both explicit judgment and endorse questions. They displayed no selectivity on ask questions, but their overall pattern of responding was selective. The exact nature of this statistical strategy is considered further in the General Discussion.

In Experiment 1, we noted that 3-year-old children did not perform as well on the explicit judgment question in the 75% versus 0% condition as they did in the other two conditions and that some children in the 3-year-old group objected to the explicit judgment question in the 75% versus 0% condition, claiming that both informants were not very good. On the basis of these observations, we speculated that the poor performance of 3-year-olds in the 75% versus 0% condition of Experiment 1 may have been due to the question wording. To test this hypothesis, we reworded the

Table 6
Number (Maximum = 2) of "Good" Responses to First and Second Explicit Judgment Questions and of "Better" Responses to Third Explicit Judgment Question

	3-year-olds ($n = 22$)		4-year-olds ($n = 30$)	
Question	*M*	*SD*	*M*	*SD*
75% correct versus 25% correct condition				
Mean number of good judgments for 75% correct (out of 2)	1.41	0.71	1.43	0.72
Mean number of good judgments for 25% correct (out of 2)	0.55	0.28	0.53	0.27
Mean number of better judgments for 75% correct (out of 2)	1.23	0.62	1.60	0.80
75% correct versus 0% correct condition				
Mean number of good judgments for 75% correct (out of 2)	1.18	0.59	1.60	0.80
Mean number of good judgments for 0% correct (out of 2)	0.86	0.43	0.47	0.24
Mean number of better judgments for 75% correct (out of 2)	1.09	0.55	1.70	0.85

Table 7
Proportion Correct and Comparisons With Chance for Experimental Measures in Experiment 2

Question	3-year-olds			4-year-olds		
	Proportion	*t*(21)	95% CI	Proportion	*t*(29)	95% CI
75% correct versus 0% correct						
Explicit judgment	.57 (.07)	1.04	.43, .71	.81 (.05)	6.04***	.71, .91
Ask	.58 (.06)	1.59	.46, .70	.72 (.05)	4.56***	.62, .82
Endorse	.55 (.05)	0.94	.45, .65	.69 (.05)	3.92**	.59, .79
75% correct versus 25% correct						
Explicit judgment	.68 (.06)	2.86*	.56, .80	.75 (.05)	5.16***	.65, .85
Ask	.49 (.06)	−0.20	.37, .61	.53 (.05)	0.61	.43, .62
Endorse	.51 (.05)	0.22	.41, .61	.62 (.05)	2.31*	.53, .72

Note. Standard errors are in parentheses. CI = confidence interval.
* $p < .05$. ** $p < .01$. *** $p < .001$.

explicit judgment questions in Experiment 2 such that it was possible for children to designate both informants as good or not very good and to differentiate between them in a third explicit judgment question.

In Experiment 2, our analysis of children's assessments of the individual informants in the 75% versus 0% condition showed that children systematically characterized one informant as good and the other as not very good. This finding suggests that explicit judgment questions in Experiment 1 were not confusing to the majority of children. In addition, the results of the 75% versus 0% condition in Experiment 2 were similar to those of Experiment 1: 4-year-olds trusted the 75% accurate informant more so than the 0% accurate informant, whereas 3-year-olds did not systematically differentiate between the two informants. Thus, it appears that 3-year-olds' poor performance in the 75% versus 0% condition of Experiment 1 was not a result of the question wording because 3-year-olds also failed this condition in Experiment 2, in which the questions were reworded.

On the basis of the results of Experiment 1, we tentatively concluded that 3-year-olds rely on an inaccuracy strategy in deciding whom to trust. That is, we hypothesized that 3-year-olds note when an informant has made at least one inaccurate claim and judge that informant to be untrustworthy. The results of Experiment 2 provide further support for this hypothesis. Three-year-olds failed to demonstrate selective trust in both the 75% versus 25% and 75% versus 0% conditions of Experiment 2.

One final unexpected result should be noted. Although 3-year-olds' performance on the explicit judgment questions was not significantly above chance in the 75% versus 0% conditions of Experiments 1 and 2, 3-year-olds' explicit judgment performance was significantly better than chance in 75% versus 25% condition of Experiment 2. This result is surprising given the more difficult nature of the 75% versus 25% task and cannot be explained by our current theoretical framework.

General Discussion

Taken together, Experiments 1 and 2 produced three main findings. First, 3-year-olds displayed selective trust but only when one of the two informants was 100% accurate. Second, there was no evidence of a relationship between selective trust and the understanding of false belief. Third, 4-year-olds performed better than 3-year-olds and were generally successful in identifying the less accurate informant across all conditions. We consider each of these findings in turn.

The pattern of results displayed by 3-year-olds is consistent with the proposal that they use an inaccuracy strategy. More specifically, we may assume that 3-year-olds operate with a simple binary coding system. At the outset, they categorize both informants as trustworthy by default. However, this default categorization is replaced by untrustworthy if an informant makes an error. Thus, 3-year-olds treat informants who make no errors as trustworthy, but they treat informants who make one or more errors as untrustworthy. On this hypothesis, we would expect 3-year-olds to display selective trust in the two conditions of Experiment 1 in which one informant was fully accurate (100% vs. 0% and 100% vs. 25%) but to respond indiscriminately in the third condition of Experiment 1 in which both informants made at least one error (75% vs. 0%). For the same reason, we would expect 3-year-olds to respond indiscriminately in both conditions of Experiment 2 (75% vs. 0% and 75% vs. 25%). This corresponds to the pattern of results observed.

As a further check on 3-year-olds' use of an inaccuracy strategy, their proportional scores in each experiment (collapsed across test questions) were analyzed by means of a one-way ANOVA, with condition as the within-subject variable. Experiment 1 yielded a main effect of condition (100% vs. 0%: $M = 0.69$, $SD = 0.21$; 75% vs. 0%: $M = 0.54$, $SD = 0.25$; 100% vs. 25%: $M = 0.73$, $SD = 0.18$), $F(2, 36) = 10.41$, $p < .001$, $\eta^2 = .37$, consistent with 3-year-olds' display of selective trust in two of the three conditions (100% vs. 0% and 100% vs. 25%). By contrast, Experiment 2 yielded no effect of condition (75% vs. 25%: $M = 0.58$, $SD = 0.19$; 75% vs. 0%: $M = 0.56$, $SD = 0.19$; $F(1, 32) = 0.67$, *ns*, consistent with their lack of selective trust in either condition.

Turning to the relationship between children's selective trust and their understanding of false beliefs, Experiment 1 showed that children's understanding of false beliefs did not predict performance on the selective trust task, and many children who failed the false belief task went on to display selective trust. Moreover, as

just noted, 3-year-olds performed above chance in two of the three conditions of Experiment 1. Yet, 3-year-olds performed systematically below chance on the false belief task.

Taken together, these findings imply that epistemic trust may not be closely tied to children's developing theory of mind. However, it is important to emphasize three caveats. First, our false belief task was quite difficult. Four-year-olds typically succeed on false belief tasks, but the 4-year-olds in Experiment 1 performed at chance level. Second, standard false belief tasks have been criticized for their information processing demands (Bloom & German, 2000). If children failed the false belief task because of information processing demands, then the conclusion that children's understanding of false belief and selective trust are unrelated would be premature. This possibility could be explored in future research by examining the relationship between selective trust and a wider range of theory-of-mind tasks. For example, Wellman and Liu (2004) have reported on a five-task scale appropriate for testing theory-of-mind development in preschool children. Third, our conclusion that false belief understanding and selective trust may be unrelated is weakened by the recent findings of DiYanni and Kelemen (2007). They found a relationship between a similar trust task and a pencil-and-paper version of the unexpected contents theory-of-mind task, controlling for age. In sum, further research is needed before researchers can be sure that there is no relationship between children's emerging understanding of belief and their selective trust in particular informants.

Finally, we may consider the performance of 4-year-olds. First, the results of both experiments confirm and extend those of Koenig and Harris (2005) and Birch et al. (2005), which also showed that selective epistemic trust improves between the ages of 3 and 4 years. Second, as noted earlier, 4-year-olds' above-chance performance in all four of the different conditions examined across the two experiments is consistent with their use of a statistical strategy. More specifically, it is feasible that 4-year-olds succeeded in all four conditions by tracking not simply whether an error had been produced but the raw frequency of errors produced by each informant. We may assume that 4-year-olds resemble 3-year-olds in their initial approach to an informant: They judge both informants as trustworthy by default. Thereafter, however, they modify that initial judgment as a function of the successive errors produced by each informant. Suppose, for example, that 4-year-olds assign each informant a default trustworthiness score of 0 and then reduce this score by 1 each time an informant makes an error. In the 100% versus 0% condition, the accurate informant would retain the initial default trustworthiness score of 0, but the inaccurate informant would end up with a trustworthiness score of −4. More generally, by tracking the raw frequency of each informant's errors, 4-year-olds would generate trustworthiness scores that differed for the two informants by at least 2 points in all four conditions. Hence, they should respond selectively in each condition, as observed.

When strictly interpreted, such a raw frequency strategy could imply not just that 4-year-olds will display selectivity across all four conditions but also that the degree of selectivity should vary with condition, depending on how far the two informants differ in the number of errors they make. To examine this possibility, we analyzed 4-year-olds' proportional scores in each experiment (collapsed across test questions) by means of a one-way ANOVA, with condition as the within-subject variable. Experiment 1 failed to yield a main effect of condition (100% vs. 0%: $M = 0.86$, $SD = 0.14$; 75% vs. 0%: $M = 0.81$, $SD = 0.22$; 100% vs. 25%: $M = 0.86$, $SD = 0.18$), $F(2, 36) = 0.90$, *ns*. Such an effect might have been expected because the less accurate informant made four more errors than did the more accurate informant in the 100% versus 0% condition but only three more errors in the other two conditions. On the other hand, Experiment 2 did yield an effect of condition (75% vs. 25%: $M = 0.65$, $SD = 0.20$; 75% vs. 0%: $M = 0.75$, $SD = 0.19$), $F(1, 32) = 13.81$, $p < .001$, $\eta^2 = .30$, consistent with the prediction that selective trust would be greater in the 75% versus 0% condition in which the less accurate informant made three more errors than did the more accurate opponent, as opposed to only two more errors in the 75% versus 25% condition. Arguably, in adopting a raw frequency strategy, 4-year-olds have difficulty in discriminating between informants and in showing selective trust, whenever the two informants differ by only a small number of errors. Thus, when informants differ by only one or two errors, 4-year-olds regard each informant as equally (un)trustworthy. On the other hand, when the two informants differ by three or four errors, 4-year-olds trust the more accurate informant. This would explain why an effect of condition emerged in Experiment 2 but not in Experiment 1. In summary, the model that we have proposed, which assumes that 4-year-olds monitor the raw frequency with which informants make errors, captures the overall success of 4-year-olds' across all four conditions and goes some way toward explaining variation in performance across those conditions.

It is worth emphasizing, however, that the findings we have reported for 4-year-olds are also consistent with a more sophisticated type of statistical strategy. Thus, 4-year-olds might monitor the proportion of inaccurate to accurate claims rather than the raw frequency of inaccurate claims. Such a strategy would lead to selective trust in the four conditions examined in Experiments 1 and 2. It would also lead children to continue to trust an informant who made several mistakes provided he or she was predominantly accurate. In future research, it should be possible to tease apart such a proportional strategy from a strategy based on raw inaccuracy (or accuracy). Suppose that 4-year-olds view three informants: A, who is accurate on four of four trials; B, who is accurate on four of eight trials; and C, who is accurate on zero of four trials. Children using a proportional strategy should prefer to trust Informant A over B and B over C. Children who attend to only the raw frequency of inaccuracy should trust A over B but not B over C. Finally, children who attend to only the raw frequency of accuracy should fail to discriminate A and B but trust B over C. Note that even if 4-year-olds fail to display such a strategy, it might be adopted by older children, granted the likely importance of maintaining trust despite occasional error. Hence, in future research it will be informative to expand the age range that is examined and the range of conditions in which children are assessed.

It is important to consider a limitation of this study and several previous studies. Both age groups showed some sensitivity to the differential accuracy of the two informants, but children's interpretation of the errors that they witnessed is unclear. They may have construed errors as playful mistakes, as deliberate lies, or as sincere but misinformed claims. Pending further research on exactly how children interpret informants' errors, several considerations may be weighed against these concerns. First, in a recent unpublished study, 3- and 4-year-olds were given four accuracy trials in which one informant was 100% accurate whereas the other informant was 0% accurate. Children were then tested for their selective trust immedi-

ately, 1 day later, and 1 week later. Despite a slight attenuation in selective trust over time, both age groups displayed greater trust in the accurate informant (as indexed by their responses to ask and endorse probes) at all three time points (Corriveau & Harris, 2006). Children's continuing selectivity after a week-long interval (during which they had no opportunity to interact with the two informants) makes it unlikely that they regarded the inaccurate informant as having initially produced only playful mistakes. Second, children's understanding and production of lies changes considerably between 3 and 4 years (Ruffman, Olson, Ash, & Keenan, 1993; Sodian, Taylor, Harris, & Perner, 1991; Talwar & Lee, 2002). Yet 3- and 4-year-olds displayed a similar pattern of selective trust across all three time periods. Hence, it is conceivable but improbable that children in both age groups regarded the inaccurate informant as producing deliberate lies. A more plausible conclusion is that children came to think of the two informants as differing in information. Consistent with this interpretation, Jaswal and Neely (2006) have recently shown that children initially trust adult informants over child informants but that preference reverses if children are given accuracy trials in which the child informant proves better informed than the adult informant.

Finally, in reflecting on the broader implications of the findings, it is worth entertaining the possibility that young children—and adults—place such a premium on accurate information that they mistrust an inaccurate informant even when the exact reasons for his or her inaccuracy remain indeterminate. By way of illustration, consider the case of Tony Blair. It remains unclear whether in the weeks leading up to the invasion of Iraq in 2003, he misled the British public about the possible existence of weapons of mass destruction or whether he was himself misled by faulty intelligence. Despite that persisting ambiguity as to whether his claims were lies or sincere errors, his reputation for trustworthiness was severely undermined when those claims proved inaccurate in that no such weapons were ever found.

In conclusion, the results from these two experiments suggest that 3- and 4-year-olds are alert to whether someone has made a false claim. They remember such errors; they subsequently judge that the person is not good at answering questions; and when faced with conflicting claims, they avoid seeking and accepting information from such an informant. However, 4-year-olds demonstrate selective trust more systematically than do 3-year-olds. Their pattern of performance suggests that, unlike 3-year-olds, 4-year-olds are using a statistical strategy—they track the number of errors, or conceivably the proportion of errors, made by each informant. Finally, the results from Experiment 1 indicate that false belief understanding and epistemic trust may be unrelated abilities that each develops between the ages of 3 and 4, but future research must be conducted to confirm this finding.

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9,* 321–324.

Birch, S., Luca, S., Frampton, K., Vauthier, S., & Bloom, P. (2005, April). *Children's assessments of what others know: The effects on learning*. Poster session presented at the Biannual Society for Research in Child Development, Atlanta, GA.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77,* B25–B31.

Clément, F., Koenig, M., & Harris, P. L. (2004). The ontogenesis of trust. *Mind and Language, 19,* 360–379.

Corriveau, K., & Harris, P. L. (2006). *Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information*. Manuscript submitted for publication.

DiYanni, C., & Kelemen, D. (2007). [Using a bad tool with good intention: Young children's imitation of adults' questionable choices]. Unpublished raw data.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5, Serial No. 242).

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. *Child Development, 59,* 26–37.

Harris, P. L., & Koenig, M. (2006). Trust in testimony: How children learn about science and religion. *Child Development, 77,* 505–524.

Harris, P. L., & Koenig, M. A. (in press). Imagination and testimony in cognitive development: The cautious disciple? In I. Roth (Ed.), *Imaginative minds.* Oxford, England: Oxford University Press.

Hogrefe, G. J., Wimmer, H. J., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development, 57,* 567–582.

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science, 17,* 757–758.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83,* B35–B42.

Koenig, M. A., Clément, F., & Harris, P. A. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science, 15,* 694–698.

Koenig, M. A., & Harris, P. A. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76,* 1261–1277.

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

Perner, J., & Wimmer, H. (1988). Misinformation and unexpected change: Testing the development of epistemic state attributions. *Psychological Research, 50,* 191–197.

Ruffman, T. R., Olson, D. R., Ash, T., & Keenan, T. (1993). The ABCs of deception: Do young children understand deception in the same way as adults? *Developmental Psychology, 29,* 74–87.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928.

Sodian, B., Taylor, C., Harris, P. L., & Perner, J. (1991). Early deception and the child's theory of mind: False trails and genuine markers. *Child Development, 62,* 468–483.

Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development, 26,* 436–444.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72,* 655–684.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75,* 523–541.

Received June 8, 2006
Revision received January 9, 2007
Accepted January 22, 2007 ■