Young Children Detect and Avoid Logically Inconsistent Sources: The Importance of Communicative Context and Executive Function

Sabine Doebel
University of Colorado

Shaina F. Rowell
University of Virginia

Melissa A. Koenig
University of Minnesota

The reported research tested the hypothesis that young children detect logical inconsistency in communicative contexts that support the evaluation of speakers' epistemic reliability. In two experiments (*N* = 194), 3- to 5-year-olds were presented with two speakers who expressed logically consistent or inconsistent claims. Three-year-olds failed to detect inconsistencies (Experiment 1), 4-year-olds detected inconsistencies when expressed by human speakers but not when read from books, and 5-year-olds detected inconsistencies in both contexts (Experiment 2). In both experiments, children demonstrated skepticism toward testimony from previously inconsistent sources. Executive function and working memory each predicted inconsistency detection. These findings indicate logical inconsistency understanding emerges in early childhood, is supported by social and domain general cognitive skills, and plays a role in adaptive learning from testimony.

Understanding that something cannot be both true and false at the same time is foundational to rationality. How and when do we develop this understanding? As adults we take for granted that contradictory statements cannot be true at the same time, a basic principle of logic first formally articulated by Aristotle (1011b: 13–14). However, a long tradition of research in developmental psychology suggests that this understanding emerges only gradually in childhood. Piaget's early investigations led him to conclude that children younger than 8 years of age could not recognize contradictions. He found children often explained physical events in ways that seemed illogical, for example, stating that a candle sinks because it is round, and a ball floats also because it is round (Inhelder & Piaget, 1958). On Piaget's view, children need to construct a concept of necessity before they can recognize inconsistencies in the world and understand, for example, that it is necessarily false that different objects behave differently because of one and the same property. Osherson and Markman (1975) conducted the first experimental investigation of children's understanding of logical inconsistency and similarly found that children younger than 8 years of age had difficulty evaluating inconsistencies, often treating them as empirical statements and indicating that they "could not tell" the truth value of statements like "The chip in my hand is blue and it is not blue."

In the time since Osherson and Markman's seminal work, several studies using more sensitive paradigms have been conducted and converge in indicating the absence of logical inconsistency understanding earlier than 6 years of age. In one study, 5-year-olds rejected the possibility that two puppets making inconsistent claims about the contents of a box could both be right (Braine & Rumain, 1981); however, this estimate was likely inflated due to the absence of a control condition in which the puppets were consistent, which would take into account correct guesses. Subsequent research that included such a control found no evidence of inconsistency understanding earlier than

6 years of age (Morris & Hasson, 2010). Similarly, children younger than 6 could not accurately judge when empirical investigation was necessary to evaluate logically inconsistent claims (Russell & Haworth, 1987). Perhaps the most compelling evidence that children younger than 6 years of age do not understand logical inconsistency was provided in a series of experiments in which children were presented with storybook characters who expressed consistent and inconsistent claims and were asked to indicate which story (or character) did not make sense (Ruffman, 1999). Five experiments found that 5-year-olds were poor at detecting inconsistencies expressed within or across two individuals, and suggest that children's difficulty was not likely due to how the test question was asked (children also performed poorly when asked which character or book was silly) or to challenges inherent in remembering, coordinating, or processing the sets of claims (Ruffman, 1999).

Thus, the extant literature strongly suggests that children younger than 6 years of age cannot detect logical inconsistency; however, this conclusion stands in contrast to more recent findings that young children monitor the quality of arguments made by sources in contexts that engage consideration of epistemic reliability. For example, 3- and 5-year-old children avoid learning from speakers who have displayed circular reasoning (Corriveau & Kurkul, 2014; see also Baum, Danovitch, & Keil, 2008; Mercier, Bernard, & Clément, 2014), and 4- and 5-year-old children preferentially endorse individuals who use logical connectives like "because" when providing testimony (Bernard, Mercier, & Clément, 2012). Moreover, 3- and 4-year-old children prefer to learn from speakers who provide compelling epistemic reasons for their beliefs (e.g., by citing perceptual evidence rather than personal desires; Koenig, 2012). These more recent findings suggest that children may be able to detect inconsistency among claims at a younger age than previously found, provided the claims are presented in a communicative context that supports the evaluation of speakers as sources of information.

Along these lines, it has been theorized that reasoning evolved to facilitate social communication by helping individuals produce and evaluate arguments (Mercier & Sperber, 2011). On this account, in social communication, "senders" put forth arguments with the goal of persuading others, which puts "receivers" in a position to evaluate those arguments. Thus, reasoning evolved, in part, to allow receivers to evaluate what they are told in order to avoid being misled. On this view, children

should reason better in communicative contexts that engage their skills in evaluating human speakers' reliability (i.e., contexts in which human speakers are offering testimony to others) than in noncommunicative contexts (Mercier, 2011).

The current research revisits the question of how and when children understand logical inconsistency by testing much younger than 6 years of age can detect them when they are presented in a communicative context that is expected to engage children's propensity to evaluate epistemic reliability. In Experiment 1, we presented preschool-aged children with two human speakers, one consistent and one inconsistent, and we asked them which of the speakers did not make sense. Unlike previous research (e.g., Ruffman, 1999), we presented children with two human informants who provided contrasting testimony (ostensibly about real events) to a third person who was seeking information. Our goal was to create a communicative context in which children would be prompted to consider the reliability of each speaker, each of whom presented their claims as true to a third party listener. In Experiment 2, we further tested whether children would be more likely to detect inconsistencies in a testimonial communicative context by comparing children's performance in a condition using the paradigm from Experiment 1 (i.e., two human speakers providing contrasting testimony to another person) to one in which children instead heard the same information read from books. We expected children to show poorer inconsistency detection when inconsistencies were read from books because the claims were not being offered in a testimonial exchange. In Experiment 2, we also explored a potential role for inconsistency detection in childhood by testing whether children would consider a source's history of inconsistency when deciding whether to trust them for new information, such as novel object labels (Experiment 1) or novel facts (Experiment 2). Finally, to gain new insight into factors supporting the development of inconsistency understanding, in Experiment 2 we tested possible predictors of improvement in inconsistency detection, with a specific interest in whether executive function might predict inconsistency detection, over and above age, working memory, and verbal knowledge.

Experiment 1

We first examined 3- to 5-year-old children's ability to detect inconsistencies in statements made by human speakers to a third party and tested

whether children would consider a speaker's prior inconsistency when given the opportunity to learn from their testimony. Children were presented with scenarios in which two human speakers were approached by a third individual for information. Across four trials, one speaker always expressed inconsistent claims and the other speaker always expressed consistent claims, and children were asked to identify which of the two speakers "did not make sense." Children were then given the opportunity to learn new information (object labels) from these speakers. We predicted that young children would show evidence of inconsistency understanding by (a) detecting inconsistent claims across two blocks of trials and (b) avoiding inconsistent speakers when learning new information.

Method

Participants

Seventy-four children participated in Experiment 1: twenty-seven 3-year-olds (M_{age} = 42.70 months, range = 36–47 months; 14 boys), twenty-four 4-year-olds (M_{age} = 53.13 months, range = 49–58 months; 12 boys), and twenty-four 5-year-olds (M_{age} = 64.83 months, range = 60–72 months; 13 boys). Children were recruited to participate from a university-maintained database of families living in a Midwestern city who previously indicated willingness to be contacted to participate in research. The sample was predominately middle to upper-middle class, university educated, and Caucasian (90% and above). Five additional children were excluded from the experiment because of experimenter error (n = 1) and uncooperativeness (n = 4). Data were collected from July 2010 through November 2010.

Materials

Ten short video clips were presented to children on a laptop. In each clip, a central female speaker approached two female speakers seated at a table (one wearing a blue shirt and the other wearing a yellow shirt) and solicited information from them in turn. When the statement expressed by the speakers involved a present object (e.g., "The frog in the box is. . ."), the object was depicted (e.g., a box was set on the table in front of the speakers with a lid on it). On selective trust test trials, novel objects (e.g., a colorful woven object and a black plastic object) were placed on the table by the central actor. Inconsistencies involved antonyms that

were expected to be familiar to young children (e.g., up/down, full/empty, big/small, loud/quiet; Murphy & Jones, 2008) and that also included wording that blocked alternative plausible interpretations (i.e., inconsistent claims that would not be easily interpreted as consistent, and vice versa; see Appendix A for the complete list of statements used in the experiment).

Procedure

The experiment was composed of three phases: (a) first block of inconsistency judgment trials, (b) selective trust test trials, and (c) second block of inconsistency judgment trials. Children were presented with a split screen image of the two main speakers and the experimenter said, "These are the people in the movie you're about to see. The girl in the yellow shirt is Kate and the girl in the blue shirt is Julie." At the beginning of the trial, the experimenter paused the video on a still frame of the two main speakers seated at a table and a central speaker standing behind and said,

> Now, I'm going to show you a movie. You will see that one of these two people, the girl in the yellow or the girl in the blue, says things that make sense and one of them says things that are wrong, that do not make sense. I want you to listen very carefully and remember what each girl says. Then I'm going to ask you who did not make sense. I want you to point to the person who did not make sense. It will be one of these two people. Okay?

Inconsistency judgment trials: first block. Prior to each test trial, the experimenter provided information to engage children's attention (e.g., "Okay, now they're going to tell us about a ball they saw today. Let's watch."). In each clip, a central speaker approached the table and asked one of the two seated speakers for information and received a response from each one in turn. For example, the central speaker said, "Can you tell me about the ball you saw today?" and the addressed speaker replied, "Today I saw a ball that was the biggest ball ever and it was the smallest ball ever." The central speaker then turned to the other seated speaker and posed the same question and the seated speaker replied, "Today I saw a ball that was the biggest ball ever and it was the softest ball ever." Figure 1 shows still frames illustrating a single inconsistency trial. The experimenter then administered a memory check asking, for example,

Figure 1. Stills from video clips used in Experiments 1 and 2 (human speakers condition). In a single inconsistency trial, each speaker is approached and asked, in turn, for information.

"Who said she saw a ball that was the smallest ever and the biggest ever?" If the child responded correctly the experimenter said, "Yes, [the girl in the yellow shirt] said she saw a ball that was the smallest ever and the biggest ever." If the child responded incorrectly, the experimenter said, "Actually, [the girl in the yellow shirt] said she saw a ball that was the smallest ever and the biggest ever." The same memory check procedure was completed for the second statement. Following the memory check, the child was asked, "Who did not make sense?"

The same speaker made inconsistent claims on each of four trials. In each test block, four of eight sets of statements were presented in a fixed order. The following were counterbalanced between subjects: which of four statement sets was presented first, which speaker (yellow shirt or blue shirt) was in the role of the inconsistent speaker, and the order of the memory checks. Within subjects, which seated speaker was addressed first by the central speaker was counterbalanced so that the speaker who was addressed first was alternated on each trial.

Selective trust test trials. Following the first inconsistency test block, children were presented with two selective trust test trials, each comprising *Ask* and *Endorse* questions.

Ask questions. Children were shown an image of a novel object and the experimenter said, "Hmm, I wonder what this is called?" A still frame of the two speakers was presented, and the experimenter said, "I bet one of these people can tell us. Who would you like to ask?" Following the child's selection the experimenter said, "Okay, let's see what they say."

Endorse questions. Children were shown a video clip in which the central speaker approached the two seated speakers and placed the novel object in the center of the table. The central speaker asked each of the seated speakers in turn, "Can you tell me what this is called?" The two speakers provided discrepant novel labels (e.g., "It's a mogit" vs. "It's a dax."). The experimenter then paused the video and asked, "What do you think it's called, a mogit or a dax?"

Inconsistency judgment trials: second block. These trials were identical to the first block of inconsistency trials with the exception that the speakers' roles were reversed, such that the speaker who was previously inconsistent in the first set of trials was consistent, and the formerly consistent speaker was now inconsistent. The same counterbalancing procedures were used as in the first block of inconsistency trials.

Results

Preliminary analyses revealed no effects of gender, statement order, or speaker role, and no interactions with variables of interest, thus all subsequent analyses collapsed across these variables. Inconsistency judgment scores were converted to mean decimal proportions for group analyses. Uncorrected p-values are reported for all post hoc tests.

Children as young as 4 years of age detected inconsistencies, and this ability improved with age. As a group, children were above chance at detecting inconsistencies, $M = .70$, $SD = .27$, $t(74) = 6.56$, $p < .001$. We conducted a two-way mixed analysis of variance (ANOVA) with age group (3-, 4-, and 5-year-olds) as the between-subjects factor, test block (first block vs. second block) as the within-subjects factor, and mean inconsistency judgment score as the dependent variable, and found a main effect of

age, $F(2, 72) = 18.77$, $p < .001$, partial $\eta^2 = .34$, a main effect of test block, $F(1, 72) = 5.39$, $p = .023$, partial $\eta^2 = .07$, and no significant interaction between these variables, $F < 1$. Children were better at detecting inconsistencies in the second block (Block 1: $M = .64$, $SD = .38$; Block 2: $M = .76$, $SD = .31$), and 5-year-olds ($M = .90$, $SD = .20$) were better at detecting inconsistencies than 3-year-olds ($M = .52$, $SD = .17$), $t(49) = 7.31$, $p < .001$, and 4-year-olds ($M = .70$, $SD = .28$), $t(46) = 2.8$, $p < .001$. Four-year-olds were also better at detecting inconsistencies than 3-year-olds, $t(49) = 2.82$, $p = .014$ (see Figure 2). To compare each age group's performance to what would be expected by chance, we collapsed across test blocks and found that both 4- and 5-year-olds were above chance in detecting inconsistencies, $t(23) = 3.58$, $p < .002$ and

$t(23) = 9.9$, $p < .001$, whereas 3-year-olds were not, $t(26) = .72$, $p = .47$. Table 1 provides children's inconsistency judgment performance by test block and age group along with comparisons against chance.

By contrast, 3-year-olds' could recall which speaker expressed which claims, although they were not as good at this as 5-year-olds. A one-way ANOVA with age group as the between-subjects factor and average memory score across the eight memory trials as the dependent variable indicated an effect of age, $F(2, 72) = 5.29$, $p < .01$. Specifically, 5-year-olds ($M = .98$, $SD = .57$) performed more accurately than 3-year-olds ($M = .83$, $SD = .18$) and 4-year-olds ($M = .88$, $SD = .21$) on the memory checks, $t(49) = 3.82$, $p < .001$ and $t(46) = 2.24$, $p = .03$, respectively, whereas the difference between 3- and 4-year-olds was not significant, $t(49) = .90$, $p = .38$. Three-year-olds who showed above-chance memory for who expressed what claims (at least five of eight memory checks correct; $n = 21$) were no better at detecting inconsistencies ($M = .52$, $SD = .29$), $t(22) = .52$, $p = .52$.

Five-year-olds used information about a speaker's prior inconsistency to guide learning on selective trust trials. Children's responses to Ask and Endorse questions were correlated in our sample, $r(75) = .29$, $p = .013$. We first report children's performance on selective trust trials collapsing across question type, followed by the same analyses separated by question type. A one-way ANOVA with age group as the between-subjects factor and mean selective trust score as the dependent variable indicated a main effect of age group, $F(2, 72) = 4.26$, $p = .018$. Five-year-olds ($M = .66$, $SD = .21$) were more likely to prefer the testimony of the consistent speaker than both 3-year-olds ($M = .51$, $SD = .22$), $t(49) = 2.50$, $p = .016$, and 4-year-olds ($M = .48$, $SD = .26$), $t(46) = 2.66$, $p = .011$. Three- and four-year-olds did not differ in the degree to which they preferred to learn from the consistent speaker, $t(49) = .46$, $p = .65$, and neither age group was above
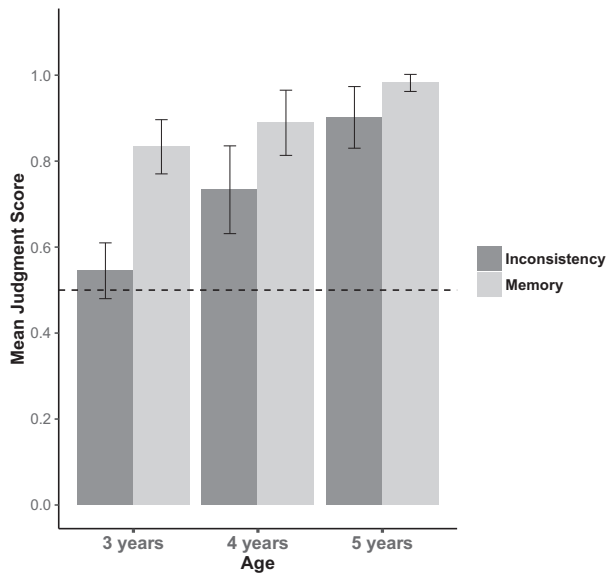

Figure 2. Children's performance in Experiment 1 on memory checks and inconsistency judgments by age. Memory performance was above chance for all age groups, and 4- and 5-year-olds also performed above chance on inconsistency judgments, but not 3-year-olds. Error bars indicate the 95% confidence intervals.

Table 1

Children's Inconsistency Detection Compared to Chance in Experiment 1

Inconsistency judgment block	3-year-olds ($n = 27$)			4-year-olds ($n = 24$)			5-year-olds ($n = 24$)		
	M	t	p	M	t	p	M	t	p
Block 1	.46 (.39)	.66	.52	.63 (.35)	1.73	.10	.88 (.26)	7.19	< .001
Block 2	.59 (.29)	1.68	.11	.78 (.35)	3.96	< .01	.93 (.16)	13.41	< .001
Total	.52 (.17)	.73	.47	.70 (.28)	3.58	< .01	.90 (.20)	9.9	< .001

Note. Mean scores indicate the decimal proportion of trials on which children correctly identified the inconsistent speaker. Standard deviations are indicated in parentheses.

chance in preferring to learn from the consistent speaker, 3-year-olds: $t(26) = .15$, $p = .89$; 4-year-olds: $t(23) = .45$, $p = .65$. Five-year-olds, on the other hand, were above chance in preferring to learn from the consistent speaker, $t(23) = 3.72$, $p = .001$.

Analysis of the data by question type revealed similar patterns. We conducted two one-way ANOVAs (corresponding to each question type) with age group as the between-subjects factor and found a main effect of age when Ask was the dependent variable, $F(2, 72) = 5.12$, $p = .008$, and a similar but somewhat weaker pattern when Endorse was the dependent variable, $F(2, 72) = 2.96$, $p = .058$. Three-year-olds were at chance on both selective trust questions, Ask: $M = .43$, $SD = .27$, $t(26) = 1.44$, $p = .16$; Endorse: $M = .59$, $SD = .34$, $t(26) = 1.41$, $p = .17$; and 4-year-olds were also at chance on both selective trust questions, Ask: $M = .52$, $SD = .23$, $t(23) = .44$, $p = .66$; Endorse: $M = .43$, $SD = .37$, $t(23) = .82$, $p = .41$. Five-year-olds, on the other hand, were above chance on both questions, Ask: $M = .65$, $SD = .23$, $t(23) = 3.08$, $p = .005$; Endorse: $M = .67$, $SD = .28$, $t(23) = 2.89$, $p = .008$. Table 2 summarizes these findings.

Discussion

In contrast to previous studies finding that children detect inconsistency no earlier than 6 years of age (e.g., Ruffman, 1999), we found evidence that children as young as 4 years of age reliably detect when a speaker expresses an inconsistent claim. In our experiment, the performance of 4-year-olds was very close to that of 5-year-olds, whereas 3-year-olds performed poorly. Our findings suggest that this is not likely due to a basic failure to remember who expressed which set of claims, consistent with previous findings (Ruffman, 1999).

Our findings are consistent with our expectation that children would be able to detect inconsistencies at younger ages in a communicative context that

engages children's propensity to evaluate the epistemic reliability of sources. Specifically, we expected that presenting children with human speakers who offered contrasting testimony to a third party would facilitate children's consideration of the content of each speaker's claims, improving their detection of problematic testimony. In Experiment 2, we further investigated this possibility by manipulating the degree to which the context in which inconsistencies were presented was likely to engage children's propensity to evaluate the reliability of testimony.

The finding that 5-year-old children used information about a speaker's history of inconsistency to guide their testimonial learning is consistent with the possibility that one function of inconsistency understanding may be to estimate a speaker's epistemic competence by evaluating the coherence of what they say (Mercier & Sperber, 2011). However, it was somewhat surprising that 4-year-old children could reliably pick out which of two speakers did not make sense but did not then avoid the speaker for new information. One possibility is that children did not avoid previously inconsistent speakers because prior inconsistencies were not seen as relevant to a new learning context involving conventional names for unfamiliar objects. If children were instead asked to learn new facts from an inconsistent source, they might be more inclined to avoid them. We tested this possibility in Experiment 2.

Our finding that 4-year-olds detected inconsistencies suggests new hypotheses about supporting skills. Specifically, between the ages of 3 and 5 years, children show dramatic improvement in executive function, the use of control over thought, emotion, and action to achieve goals (Diamond, 2013). Executive function often involves overriding habit or a dominant response tendency (Gerstadt, Hong, & Diamond, 1994; Morton & Munakata, 2002; Zelazo et al., 2003), and it is plausible that inconsistency detection might depend on executive function in nontrivial ways. First, being able to evaluate the logical compatibility of two or more claims may require one to inhibit the tendency to accept information as true and coherent. For example, children with lower levels of inhibitory control have difficulty ignoring misleading testimony (Jaswal Perez-Edgar, Kondrad, Palmquist, Cole, et al., 2014). A second possibility, consistent with ideas put forth by Osherson and Markman (1975), is that inconsistency detection requires inhibiting the default tendency to treat all claims as empirically verifiable and instead flexibly attending to their

Table 2

Children's Preferences to Learn New Information From a Previously Consistent Source

	Experiment 1			Experiment 2	
	3-year-olds	4-year-olds	5-year-olds	4-year-olds	5-year-olds
Ask	.43 (.27)	.52 (.23)	.65 (23)**	.55 (.25)	.54 (.25)
Endorse	.59 (.34)	.43 (37)	.67 (28)**	.55 (.27)	.59 (.30)*
Both	.51 (.22)	.48 (.26)	.66 (.21)***	.55 (.21)†	.57 (.21)**

Note. †$p < .10$. *$p < .05$. **$p < .01$.***$p < .001$.

logical form. Thus, in Experiment 2 we also examined whether executive function is related to logical inconsistency detection in children.

Experiment 2

In Experiment 2, we focused exclusively on 4- and 5-year-old children, the age range in which we found evidence of inconsistency understanding in Experiment 1. Experiment 2 had three aims. First, we tested the hypothesis that young children are better at detecting inconsistencies when they are presented in a communicative context that activates consideration of the source's epistemic reliability. Specifically, we presented children either with human speakers who expressed inconsistencies to a third party, as in Experiment 1, or with books that contained inconsistencies that were read aloud by an experimenter. As in Experiment 1, we expected the human speakers condition to facilitate inconsistency detection because the presence of two speakers providing contrasting testimony to a third party should engage children's propensity to evaluate the source and what they say. The book-reading context, on the other hand, was not expected to have this effect because the claims were not being staked directly by anyone or to anyone; rather, they were simply being read from a book by an experimenter. Second, we tested whether young children would be more inclined to use prior inconsistency to guide learning from testimony if the inconsistencies and the later testimony both involved factual claims. In Experiment 2, sources provided testimony about novel facts (e.g., claims about what unfamiliar animals eat and where they live) instead of names for novel objects. Third, we tested whether executive function predicts inconsistency detection by assessing children's performance on a measure of executive function after accounting for working memory and verbal knowledge (both of which are known to be associated with executive function; Cragg & Nation, 2010; Garon, Bryson, & Smith, 2008).

Method

Participants

Sixty 4-year-olds (M_{age} = 53.53 months, range = 48–59 months; 27 boys) and sixty 5-year-olds (M_{age} = 66.13, range = 60–71 months, 35 boys) were recruited from the same university-maintained database described in Experiment 1. Five additional children were excluded from the experiment due to uncooperativeness. Data were collected from February 2013 through October 2014.

Design and Materials

Children were randomly assigned to either the human speakers or book-reading condition. In the human speakers condition, children were shown the same video clips as in Experiment 1, in which two human speakers made claims (to a third person) that were consistent or inconsistent. In the book-reading condition, children were presented with two books—a yellow book and a blue book—from which the experimenter read pairs of consistent and inconsistent statements that were comparable to those expressed in the human speakers condition. The experiment consisted of four inconsistency judgment trials, eight trials composing a selective trust test phase, four additional inconsistency judgment trials in which the speaker's roles were reversed, and measures of working memory, executive function, and verbal knowledge.

Procedure

Human speakers condition. The procedure for the human speakers condition was identical to what was described in Experiment 1 with the exception of the selective trust phase (described in detail below).

Book-reading condition. Children were seated at a table on which there were two books place in front of the child. The experimenter said:

> Look at these books. We have a yellow book and a blue book. I am going to read some things from these books and you will see that one of them has things in it that make sense and one of them has things in it that are wrong, that do not make sense. I want you to listen very carefully and remember what each book says. Then I'm going to ask which one had things in it that did not make sense. Okay?

Inconsistency judgment trials: first block. Prior to the first trial, the experimenter provided information to engage the child's attention (e.g., "Okay, we're going to read what the books say about a ball."). The experimenter then opened one of the books and read a set of statements. For example, on one trial the experimenter said, "Okay, let's see what this book says. It says that someone saw a ball that was the smallest ball ever and that was the biggest ball ever at the same time." The

experimenter then closed the book and placed it on the table and opened the other book, saying, "Now let's see what this book says." The experimenter then proceeded to read a claim from that book: "It says that someone saw a ball that was the smallest ball ever and that was the softest ball ever at the same time." The content of the statements was the same as in the human speakers condition, although the phrasings were slightly modified so that they could be expressed as coming from the book without anthropomorphizing it or implying that a person was expressing the claims via the book (see Appendix B, for the complete list of the statements used in the experiment).

The experimenter then administered a memory check, asking the child to indicate which book contained which statement (e.g., "Which book talked about a ball that was the smallest ever and the biggest ever? And which book talked about a ball that was the smallest ever and the softest ever?"). After each trial, following the memory check, the experimenter asked the child, "Which book did *not* make sense?" The same book was inconsistent on each of four trials. Counterbalancing procedures were the same as in Experiment 1.

Selective trust test trials. Following the first inconsistency test phase, children were presented with four selective trust test trials in order to test whether children would selectively avoid learning from the inconsistent source in each condition. As in Experiment 1, each selective trust trial comprised described below.

Ask questions: book-reading condition. Children were shown a printed and laminated picture of a novel creature and were asked to select a book that could provide more information about it. For example, children were shown a "pangolin" and the experimenter said, "This is a pangolin. I wonder what it eats. Which book should we check?"

Endorse questions: book-reading condition. Once a selection was made, the experimenter said, "Okay, let's see what it says." The experimenter opened the book that the child selected and read out the book's response (e.g., "The yellow book says pangolins eat roots."). Then the experimenter opened the other book and provided its response (e.g., "Now let's see what this blue book says. . .The blue book says pangolins eat bugs."). The experimenter then asked the child for their judgment about which book provided the correct response (e.g., "The yellow book says pangolins eat roots and the blue book says pangolins eat bugs. What do you think pangolins eat, roots or bugs?"). The dependent variable derived from the selective trust phase was the mean proportion of responses to the Ask and Endorse questions in which children preferred to learn from the consistent versus inconsistent book.

Ask and Endorse questions: human speakers condition. In this condition the selective trust trials mirrored those presented in the book-reading condition, with the exception that the wording was adjusted to reflect that a human speaker was expressing the claims. To ensure that the conditions were closely matched, children in the human speakers condition were also presented with the same laminated pictures of the novel creatures, and the experimenter reported each speaker's testimony. For example, children were shown a "pangolin" and the experimenter said, "This is a pangolin. I wonder what it eats. I bet one of these people can tell us. Who would you like to ask?" To assess children's endorsement, the experimenter then said (for example):

Now let's see what they say. The girl in yellow says pangolins eat roots. Now let's see what the other girl says. The girl in blue says pangolins eat bugs. She [E points to girl in yellow] says pangolins eat roots, and she [E points to girl in blue] says pangolins eat bugs. What do you think Pangolins eat, roots or bugs?

Unlike Experiment 1, the order in which the speakers gave their testimony was dependent on children's responses to the Ask questions. This modification made the procedure more comparable to the procedure in the book-reading condition in which it was more naturalistic to open the book that the child selected.

Inconsistency judgment trials: second block. As in Experiment 1, these four test trials were identical to the first block of inconsistency trials except that the previously consistent source in each condition (book-reading or human speakers) was now always inconsistent, and vice versa. In addition, as in Experiment 1, the four sets of statements used in the first or second phase were also counterbalanced.

Forward and backward digit span. Forward and backward digit span tasks were administered to assess working memory and executive function, respectively (Carlson, 2005; Davis & Pratt, 1995). In the forward digit span task, children had to repeat strings of digits articulated by the experimenter in the same order. The experimenter started with a two-digit string (e.g., "two, one") and increased the length (e.g., "one," "six," "four". . .) until children could not repeat them correctly. In the backward

digit span task, children were instructed to repeat the string of digits in the reverse order. Success on the backward digit span task is presumed to require manipulation of mental representations of the digits (to reorder them), and also inhibition of the tendency to repeat the numbers in the same order the experimenter articulated them. The tasks were administered in a fixed order, with the forward digit span administered first, which was expected to enhance the prepotency of the forward span rule and increase demands on executive function during the backward digit span task.

Peabody Picture Vocabulary Test. The Peabody Picture Vocabulary Test (PPVT) is a measure of receptive vocabulary and was included as an index of verbal knowledge (Dunn & Dunn, 2007). On each item children were shown four pictures and provided with a word that corresponded to one of the pictures and were asked to point the picture that best corresponds to the word. Children's scores reflected the number of items to which they responded correctly minus their total errors.

Results

Preliminary analyses revealed no effects of gender, statement order, or source role (i.e., which book or speaker was consistent vs. inconsistent), and no interactions with the variables of interest, thus all subsequent analyses collapsed across these variables. As in Experiment 1, individual children's inconsistency judgment scores were averaged for group analyses.

Children were better at detecting inconsistencies when statements were presented in a testimonial communicative context. We conducted a three-way mixed ANOVA with age group (4-year-olds vs. 5-year-olds) and condition (human speakers vs. book-reading) as between-subjects factors, inconsistency test block (first vs. second) as the within-subjects factor, and mean inconsistency score as the dependent variable. We found main effects of age group, $F(1, 116) = 7.04$, $p = .009$, partial $\eta^2 = .06$, and condition, $F(1, 116) = 5.72$, $p = .018$, partial $\eta^2 = .05$. The main effects of age and condition were qualified by a significant interaction between these factors, $F(1, 116) = 4.83$, $p = .003$, partial $\eta^2 = .07$ (see Figure 3). To interpret the interaction, we examined children's inconsistency judgments by age group and condition, collapsing across test block. Four-year-old children in the human speakers condition ($M = .75$, $SD = .21$) were above chance in detecting inconsistencies, $t(29) = 6.67$, $p < .001$, and were significantly better at detecting inconsistencies than 4-
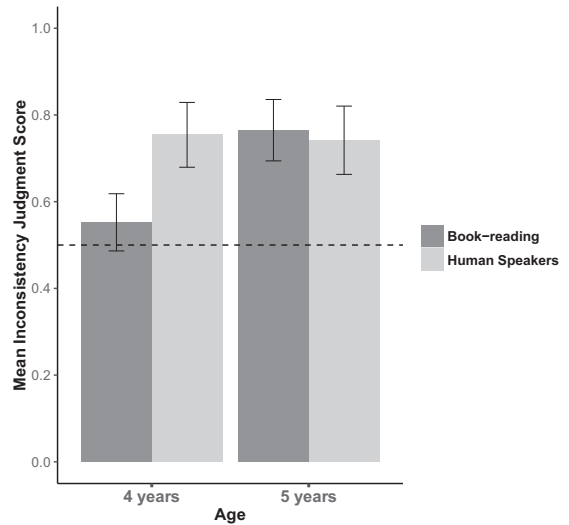

Figure 3. Children's performance in Experiment 2 on inconsistency judgments by condition. Five-year-olds performed well in both conditions, but 4-year-olds were above chance only in the human speakers condition. Error bars indicate the 95% confidence intervals.

year-olds in the book-reading condition ($M = .55$, $SD = .18$), $t(58) = 3.96$, $p < .001$, who did not differ from chance, $t(29) = .92$, $p = .36$. As illustrated in Figure 3, 5-year-olds were above chance in both the human speakers and book-reading conditions, $t(29) = 4.47$ and $t(29) = 3.98$, $ps < .001$, respectively, and in contrast to the 4-year-olds, there was no significant difference in inconsistency detection between the human speakers condition ($M = .74$, $SD = .22$) and book-reading condition ($M = .76$, $SD = .20$), $t(58) = 0.37$, $p = .71$. Table 3 shows rates of inconsistency detection by age, block, and condition, along with uncorrected comparisons to chance.

Although children showed better memory with age for which source expressed which claims, memory performance did not vary by condition. We conducted a two-way ANOVA with age group (4-year-olds vs. 5-year-olds) and condition (human speakers vs. book-reading) as between-subjects factors, and average memory score across the eight memory check trials as the dependent variable. We found a significant main effect of age group, $F(3, 116) = 18.32$, $p < .001$, partial $\eta^2 = .14$ (see Figure 4). Post hoc analyses indicated that 5-year-olds ($M = .96$, $SD = .08$) performed better than 4-year-olds ($M = .86$, $SD = .16$) on the memory checks. Children in the human speakers condition ($M = .91$, $SD = .12$) were no better than children in the book reading condition ($M = .90$, $SD = .15$) at remembering which source expressed inconsistencies, $F < .1$.

Table 3

Children's Inconsistency Detection Compared to Chance in Experiment 2

Condition and inconsistency judgment block	4-year-olds (*n* = 60)			5-year-olds (*n* = 60)		
	M	*t*	*p*	*M*	*t*	*p*
Human speakers (*n* = 30)						
Block 1	.75 (.26)	5.21	< .001	.71 (.29)	3.98	< .001
Block 2	.76 (.32)	4.45	< .001	.77 (.28)	5.36	< .001
Total	.75 (.21)	6.67	< .001	.74 (.22)	6.02	< .001
Book reading (*n* = 30)						
Block 1	.55 (.30)	.92	.36	.74 (.29)	4.48	< .001
Block 2	.56 (.35)	.91	.37	.79 (.25)	6.49	< .001
Total	.55 (.18)	1.56	.13	.76 (.20)	7.34	< .001

Note. Means indicate the proportion of trials on which children correctly identified the inconsistent speaker. Standard deviations are indicated in parentheses.
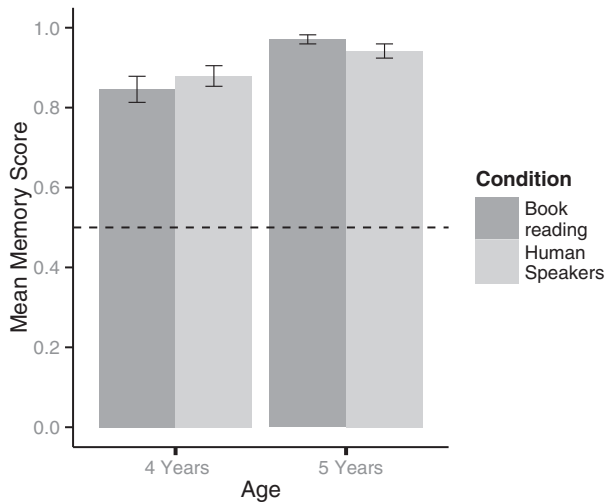

Figure 4. Children's performance in Experiment 2 on memory judgments by condition. Children in both age groups and conditions performed well above chance when answering questions about which source expressed which set of claims. Error bars indicate the standard error of the mean.

As illustrated in Figure 4, children in both age groups in both conditions were above chance in their memory for which source expressed which claims (book-reading condition: 4-year-olds: *M* = .84, *SD* = .17, *t*(29) = 10.59, *p* < .001; 5-year-olds: *M* = .97, *SD* = .06, *t*(29) = 40.93, *p* < .001; human speakers condition: 4-year-olds: *M* = .88, *SD* = .14, *t*(29) = 14.72, *p* < .001; 5-year-olds: *M* = .94, *SD* = .09, *t*(29) = 24.94, *p* < .001.

Children used information about a speaker's prior inconsistency to guide learning of new facts and this pattern did not vary by age group.

Overall, children were above chance in preferring the consistent to the inconsistent source for information about novel facts, *M* = .56, *SD* = .21, *t*(119) = 3.07, *p* = .003. To examine effects of age and condition, we conducted a two-way ANOVA with age group (4-year-olds vs. 5-year-olds) and condition (human speakers vs. book reading) as between-subjects factors, and mean selective trust score as the dependent variable. There were no main effects of age or condition, *F*s < 1, and no significant interaction between these factors, *F*(1, 116) = 2.31, *p* = .13. Exploratory analyses indicated that 4-year-olds were marginally above chance in preferring to learn new information from the consistent source (*M* = .55, *SD* = .21), *t*(59) = 1.88, *p* = .065, and 5-year-olds performed similarly (*M* = .57, *SD* = .21) and were above chance, *t*(59) = 2.45, *p* = .017.

Analyzing the data by question type revealed similar patterns. We conducted two separate two-way ANOVAs for each question type (Ask and Endorse) with age and condition as between-subjects factors. In both analyses, there were no significant main effects of age or condition, nor any effect of their interaction, all *F*s < 1.3. Children were above chance in preferring to seek information from the consistent source, *M* = .55, *SD* = .25, *t*(119) = 1.99, *p* = .049, and preferring to endorse the consistent source, *M* = .57, *SD* = .28, *t*(119) = 2.66, *p* = .009. These findings are summarized in Table 2 and are presented along with the findings from Experiment 1 for comparison.

Finally, we found evidence that executive function predicts inconsistency detection in 4- to 5-year-old children. We tested this relation using multiple regression, simultaneously entering the following predictors into the model: age, condition, backward digit span score, forward digit span score, PPVT, and a term representing the interaction between age and condition. The dependent variable was performance across the eight inconsistency judgment trials. Model diagnostics did not indicate multicollinearity (Variance Inflation Factors < 1.78, tolerances > .56) or nonlinearity of the standardized residuals. Cook's distance was used to investigate influential data points, and none were detected (all Cook's *D* values < 1). Table 4 presents the bivariate correlations and Table 5 presents a summary of the multiple regression results. As expected, forward and backward digit span scores were correlated, possibly due to shared task demands. Although the bivariate correlations between inconsistency judgment scores and each predictor variable were significant, only backward digit span and forward digit

Table 4

Correlations Among Variables in Multiple Regression Testing Predictors of Inconsistency Detection in Experiment 2

Variable	1	2	3	4	5
1. Inconsistency score	—				
2. Age	.36**	—			
3. BDS	.49**	.50**	—		
4. FDS	.42**	.44**	.53**	—	
5. PPVT	.43**	.60**	.52**	.36**	—
6. Condition[a]	.20*	.05	.03	−.06	.17

Note. BDS = backward digit span; FDS = forward digit span; PPVT = Peabody Picture Vocabulary Test. [a]Spearman's rho. *$p < .05$. **$p < .001$.

Table 5

Summary of Multiple Regression Testing Predictors of Inconsistency Detection in Experiment 2

Variable	B	SE B	β	t	p
Constant	.179	.172		1.042	.299
Age	−.001	.003	−0.003	−0.031	.975
BDS	.043	.018	0.261	2.455	.016
FDS	.054	.024	0.225	2.255	.026
PPVT	.009	.001	0.188	1.736	.086
Condition	.277	.154	1.278	1.800	.075
Age × Condition	−.004	.003	−1.153	−1.624	.108

Note. BDS = backward digit span; FDS = forward digit span; PPVT = Peabody Picture Vocabulary Test.

span independently predicted significant variance in the model. There was also a trend suggesting PPVT score might predict unique variance in inconsistency judgment performance.

Discussion

Experiment 2 extended the findings of Experiment 1 by replicating the result that children as young as 4 years of age can detect inconsistencies, and by showing that they do so if inconsistencies are expressed as testimony offered by human speakers but not if they are reported from books. We found the same pattern of good performance on memory checks as in Experiment 1, and did not find evidence that memory performance was worse in the book-reading condition, suggesting that poor performance in that condition cannot be attributed to relatively poor memory for the claims. Our findings are consistent with our hypothesis that young children's nascent reasoning abilities are more likely to be expressed in contexts that engage their consideration of a source's epistemic reliability.

We also found further evidence that inconsistency detection supports selective learning from testimony, with 4- and 5-year-olds using a source's prior inconsistency to guide their learning of new facts. Unlike in Experiment 1, we did not find a main effect of age, consistent with the possibility that once children can detect a source's inconsistency, they will use that information to guide testimonial learning.

Finally, we found new evidence that executive function, working memory, and possibly verbal knowledge, play roles in young children's inconsistency detection. This suggests that inconsistency detection depends in part on the development of domain general processes.

General Discussion

The reported experiments provide new insights concerning how and when logical inconsistency understanding emerges in childhood, and suggest a role for it in social learning. Previous research indicates that children younger than 6 years of age cannot detect logical inconsistency, however our findings across two experiments indicate that even 4-year-olds can do so if the claims they are evaluating are presented in a communicative context that encourages consideration of the source's epistemic reliability. Moreover, children remembered when sources were inconsistent and avoided learning new information from them, suggesting that logical reasoning in early childhood supports children's social learning. Finally, we found evidence that executive function, working memory, and possibly verbal knowledge each independently contribute to the emergence of inconsistency understanding during childhood.

Our finding that children as young as 4 years of age were able to detect inconsistencies when human speakers provided testimony is consistent with our expectation that such a context engages children's propensity to evaluate the epistemic reliability of informants. When considering whether a source is reliable or not, children may bring to bear whatever skills they have to make that assessment, including nascent logical reasoning skills. Put another way, the testimonial context may serve to prompt an epistemically vigilant stance, and as a result children may evaluate arguments and claims more carefully than they would otherwise, given their dependency on others for information and their risk of being misled (Mascaro & Sperber, 2009; Mercier & Sperber, 2011). Consistent with this interpretation, children also avoided learning new information from previously inconsistent sources.

The human speakers context may have heightened consideration of epistemic reliability because children found the speakers more engaging and worth attending to than the experimenter reading statements from books. Consistent with evidence that social interaction serves a gating function for computational learning mechanisms (e.g., Kuhl, 2004; Kuhl, Tsao, & Liu, 2003), it may be that observing speakers who were engaged in an active, social, testimonial exchange captured children's attention and enhanced their processing of logical connectives and the predicates they connect. On the other hand, children showed comparable memory for claims across conditions, and the book-reading condition also had social elements (i.e., the experimenter read to the child with a positive demeanor in a child-directed register) that would also be expected to engage children's interest and attention. We suggest that human speakers providing testimony to another was nevertheless more engaging for children than testimony being read from books, leading children to more deeply evaluate and process the information conveyed.

It is also possible that testimony that is read from a book is more likely to be accepted by younger children than testimony that is simply spoken. That is, the younger children in Experiment 2 may have approached the book-reading context with relatively low concern about epistemic reliability, as they come to treat books as sources of information about the world (Ganea, Ma, & DeLoache, 2011; Ganea, Pickard, & DeLoache, 2008). This would be consistent with research indicating that early readers are more likely to accept unexpected suggestions (calling a dog-like thing a "cat") when conveyed via print versus orally (Eyden, Robinson, Einav, & Jaswal, 2013). Future research can further examine and clarify how different contexts and presentation formats influence children's detection of problematic testimony.

Although children may be better at detecting logical inconsistencies when they are provided in a communicative context that engages children's consideration of epistemic reliability, this does not mean that younger children cannot under any circumstance detect problems with testimony provided by nonhuman informants. For example, children as young as 3 years of age will reject information from technological sources (i.e., computers) that previously provided inaccurate information (e.g., displaying the color orange in response to a question about the color of grass) (Danovitch & Alzahabi, 2013). Detecting inconsistent sources is likely more challenging than detecting inaccurate ones for a variety of reasons, including that inconsistency detection operates on more than one proposition maintained in mind and thus likely depends to a greater extent on general cognitive processes like working memory, executive function, and verbal skills. As such, it may be that a communicative context serves to facilitate the critical evaluation of claims in those children who are just developing the cognitive skills to detect inconsistencies, such as the 4-year-olds in our study. Five-year-olds, given their cognitive and linguistic abilities, may not derive the same benefits from contextual support, as indicated by their similar performance in inconsistency detection across the human speakers and book-reading conditions.

The finding in both experiments that children avoided learning new information from previously inconsistent sources indicates that one function of logical reasoning in early childhood may be to evaluate a source's epistemic reliability (Mercier & Sperber, 2011). On the other hand, our findings do not reveal what it is about inconsistency that children find problematic. Children may grasp that inconsistencies do not make sense and are uninformative, as indicated by their explicit judgments and selective learning, but they may not find them *illogical* or indicative of irrationality (see also Shtulman & Carey, 2007). The extent of young children's appreciation of the nature of logical inconsistency is a key question for future research.

What causes children's difficulty detecting inconsistencies? Although our findings indicate that children as young as 4 years of age can detect logical inconsistencies when presented in a facilitating context, many 4-year-olds did not detect inconsistencies and there was still room for improvement among 5-year-olds. One proposal is that children have "incomplete insight into logical necessity" due to an inability to evaluate the mental models of reality they construct (Ruffman, 1999; see also Morris & Hasson, 2010, for a similar view). By contrast, others have suggested that children have difficulty identifying inconsistencies because they lack the ability to reflect on language and to consider linguistic expressions independently of reality (Osherson & Markman, 1975). On this view, children mistakenly view logical inconsistencies as empirical claims to be evaluated by investigating how things stand in the world. Our findings are consistent with both of these views. Consistent with the suggestions of Osherson and Markman (1975), our findings suggest that executive function may play a role in helping children override the tendency to view inconsistencies as empirical claims and

evaluate these claims independently of reality. In addition, it is also possible that executive function and working memory play roles in the development of a concept of logical necessity, by allowing one to evaluate and compare propositions that are maintained in mind. Thus, consistent with Ruffman's (1999) view, developments in executive function and working memory could allow children to both construct and evaluate mental models.

Although verbal knowledge was not a significant predictor of inconsistency detection, we did find a statistical trend, consistent with previous findings that having older siblings (and thus greater opportunities for verbal skill development) is associated with inconsistency detection (Ruffman, 1999). There are a number of reasons why verbal knowledge could facilitate inconsistency understanding. As we suggested earlier, it may be that having certain conceptual knowledge is key to being able to evaluate the relation between inconsistent claims (e.g., knowledge of what it means to be a certain age supports the understanding that one cannot be both 3 and 7 years old). Another possibility is that verbal knowledge predicts understanding of logical terms such as "and" and "not." Children's early conversational experience involving such words is largely informal and rarely are they presented in logical contexts (Morris, 2008). Future research can examine whether children who receive more experience with these words in logical contexts are better at detecting logical inconsistencies.

Conclusion

The reported findings provide new insights about how and when logical inconsistency understanding first emerges and suggests that it supports social learning by helping children evaluate informants and what they say. More broadly, the findings suggest that young children may reason better with increases in executive function and when information is presented in a communicative context that engages their skills in evaluating human speakers.

References

Aristotle. *Metaphysics*. Book 4. Section 1011b.

Baum, L. A., Danovitch, J. H., & Keil, F. C. (2008). Children's sensitivity to circular explanations. *Journal of Experimental Child Psychology*, *100*, 146–155. doi:10.1016/j.jecp.2007.10.007

Bernard, S., Mercier, H., & Clément, F. (2012). The power of well-connected arguments: Early sensitivity to the connective because. *Journal of Experimental Child Psychology*, *111*, 128–135. doi:10.1016/j.jecp.2011.07.003

Braine, M. D. S., & Rumain, B. (1981). Development of comprehension of "or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, *31*, 46–70. doi:10.1016/0022-0965(81)90003-5

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595–616. doi:10.1207/s15326942dn2802_3.

Corriveau, K. H., & Kurkul, K. E. (2014). "Why does rain fall?": Children prefer to learn from an informant who uses noncircular explanations. *Child Development*, *85*, 1827–1835. doi:10.1111/cdev.12240

Cragg, L., & Nation, K. (2010). Language and the development of cognitive control. *Topics in Cognitive Sciences*, *2*, 631–642. doi:10.1111/j.1756-8765.2009.01080.x

Danovitch, J. H., & Alzahabi, R. (2013). Children show selective trust in technological informants. *Journal of Cognition and Development*, *14*, 499–513. doi:10.1080/15248372.2012.689391

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology*, *47*, 25–31. doi:10.1080/00049539508258765

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. doi:10.1146/annurev-psych-113011-143750

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test (PPVT–4)*. Upper Saddle River, NJ: Pearson Education.

Eyden, J., Robinson, E. J., Einav, S., & Jaswal, V. K. (2013). The power of print: Children's trust in unexpected printed suggestions. *Journal of Experimental Child Psychology*, *116*, 593–608. doi:10.1016/j.jecp.2013.06.012

Ganea, P. A., Ma, L., & DeLoache, J. S. (2011). Young children's learning and transfer of biological information from picture books to real animals. *Child Development*, *82*, 1421–1433. doi:10.1111/j.1467-8624.2011.01612.x

Ganea, P. A., Pickard, M. B., & DeLoache, J. S. (2008). Transfer between picture books and the real world by very young children. *Journal of Cognition and Development*, *9*, 46–66. doi:10.1080/15248370701836592

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*, 31–60. doi:10.1037/0033-2909.134.1.31

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children 3 1/2 – 7 years old on a stroop-like day-night test. *Cognition*, *53*, 129–153. doi:10.1016/0010-0277(94)90068-X.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York, NY: Basic Books.

Jaswal, V. K., Pérez-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*, 965–976. doi:10.1111/desc.12187

Koenig, M. A. (2012). Beyond semantic accuracy: Preschoolers evaluate a speaker's reasons. *Child Development*, *83*, 1051–1063. doi:10.1111/j.1467-8624.2012.01742.x

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831–843. doi:10.1038/nrn1533

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 9096–9101. doi:10.1073/pnas.1532872100

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition*, *112*, 367–380. doi:10.1016/j.cognition.2009.05.012

Mercier, H. (2011). Reasoning serves argumentation in children. *Cognitive Development*, *26*, 177–191. doi:10.1016/j.cogdev.2010.12.001

Mercier, H., Bernard, S., & Clément, F. (2014). Early sensitivity to arguments: How preschoolers weight circular arguments. *Journal of Experimental Child Psychology*, *125*, 102–109. doi:10.1016/j.jecp.2013.11.011

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57–74. doi:10.1017/S0140525X10000968

Morris, B. J. (2008). Logically speaking: Evidence for item-based acquisition of the connectives AND & OR. *Journal of Cognition and Development*, *9*, 67–88. doi:10.1080/15248370701836600

Morris, B. J., & Hasson, U. (2010). Multiple sources of competence underlying the comprehension of inconsistencies: A developmental investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 277–287. doi:10.1037/a0017519

Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, *40*, 255–265. doi:10.1002/dev.10033

Murphy, M. L., & Jones, S. (2008). Antonyms in children's and child-directed speech. *First Language*, *28*, 403–430. doi:10.1177/0142723708091047

Osherson, D. N., & Markman, E. (1975). Language and the ability to evaluate contradictions and tautologies. *Cognition*, *3*, 213–226. doi:10.1016/0010-0277(74)90009-2

Ruffman, T. (1999). Children's understanding of logical inconsistency. *Child Development*, *70*, 872–886. doi:10.1111/1467-8624.00063

Russell, J., & Haworth, H. M. (1987). Perceiving the logical status of sentences. *Cognition*, *27*, 73–96. doi:10.1016/0010-0277(87)90036-9

Shtulman, A., & Carey, S. (2007). Improbable or impossible? How children reason about the possibility of extraordinary events. *Child Development*, *78*, 1015–1032. doi:10.1111/j.1467-8624.2007.01047.x

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., . . . Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*, i–151. doi:10.1111/j.1540-5834.2003.06803001.x

Appendix A

Statements Used in Inconsistency Judgment Task in Experiment 1 and Human Speakers Condition in Experiment 2

Prompt	Consistent response	Inconsistent response
Can you tell me about the ball you saw today?	Today I saw a ball that was the smallest ball ever and it was the softest ball ever at the same time.	Today I saw a ball that was the smallest ball ever and it was the biggest ball ever at the same time.
Can you tell me about your brother Ben/Max?	My brother Ben is 3 years old, and not only that, he's still little.	My brother Max is 3 years old, and not only that, he's 7 years old.
Can you tell me about the cup you saw yesterday?	Yesterday, I saw a cup that was totally full of water and it was cold at the same time.	Yesterday, I saw a cup that was totally full of water and it was empty at the same time.
Can you tell me about the frog in this box?	The frog in the box is a boy frog, and not only that, it's a small frog.	The frog in the box is a boy frog, and not only that, it's a girl frog.
Can you tell me about your friend John/Mark?	My friend Mark always walks fast and he always walks far.	My friend John always walks fast and he always walks slow.
Can you tell me about the tree you saw yesterday?	Yesterday, I saw a tree that is the tallest in the world and it's the greenest in the world.	Yesterday, I saw a tree that is the tallest in the world and it's the shortest in the world.
Can you tell me about this box?	This box is full of toys and it has a ball in it.	This box is full of toys and it is empty.
Can you tell me about the baby you saw today?	Today, there was a baby that was crying really loud and was kicking at the same time.	Today, there was a baby that was crying really loud and was quiet at the same time.

Appendix B

Statements Used in Inconsistency Judgment Task in Book-Reading Condition in Experiment 2

Prompt	Consistent response	Inconsistent response
Now we're going to read what the books say about a ball.	Now let's see what this book says. It says that someone saw a ball that was the smallest ball ever and that was the softest ball ever at the same time.	It says that someone saw a ball that was the smallest ball ever and that was the biggest ball ever at the same time.
Now we're going to read what the books say about a boy named John and a boy named Dylan.	It says there is a boy named John who is 3 years old and, not only that, he's still little.	It says that there is a boy named Dylan who is 3 years old and, not only that, he is 7 years old.
Now we're going to read what the book says about a cup.	It says someone saw a cup that was totally full of water and it was cold at the same time!	It says someone saw a cup that was totally full of water and it was empty at the same time!
Now we're going to read what the books say about a frog.	It says someone saw a frog in a box that is a boy frog, and not only that, it's a small frog.	It says someone saw a frog in a box that is a boy frog, and not only that, it's a girl frog.
Now we're going to read what the books say about a girl named Molly and a girl named Jessie.	It says that there is a girl named Molly who always walks fast and always walks far.	It says that there is a girl named Jessie who always walks fast and always walks slow.
Now we're going to read what the books say about a tree.	It says yesterday someone saw a tree that's the tallest in the world and it's the greenest in the world.	It says yesterday someone saw a tree that's the tallest in the world and it's the shortest in the world.
Now we're going to read what the books say about a box.	It says that someone saw a box that is full of toys and it has a ball in it.	It says that someone saw a box that is full of toys and it is empty.
Now we're going to read what the books say about a baby.	It says today there was a baby that was crying really loud and it was kicking at the same time.	It says today there was a baby that was crying really loud and it was quiet at the same time.