Cognitive Development 54 (2020) 100871

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Children's developing understanding that even reliable sources need to verify their claims

Lucas Payne Butler[a,*], Hailey M. Gibbs[a], Nadia S. Tavassolie[b]

[a] Department of Human Development and Quantitative Methodology, University of Maryland, College Park, United States
[b] Department of Psychology, Temple University, United States

ARTICLE INFO

Keywords:
Epistemic cognition
Social cognition
Selective trust
Pedagogy

ABSTRACT

In a society inundated with information, fostering empirical habits of mind in young children is critically important. Perhaps the most fundamental epistemic principle is that claims, no matter who is making them, must be supported by sufficient evidence. Building on past research on young children's selective social learning, three experiments with 3- to 7-year-old children (N = 136) investigated when children come to understand this central principle. The results suggest clear age differences: By age 4, children consistently selected an individual who verified their claims over one who did not and by age 6, children were able to do so even if that act of verification conflicted with that individual's past history of reliability. Further, with age, children's ability to explicitly justify those choices in terms of informants' epistemic practices improved dramatically. Overall, this work presents an important next step in investigating children's developing understanding of the process by which evidence supports epistemic claims.

1. Introduction

A pressing issue facing society is ensuring a populace that approaches the world with a critical, evaluative lens. Particularly in an era when nearly unlimited information is available at our fingertips, we must continually assess whether claims we encounter are supported by sufficient evidence. This requires the cultivation of a set of empirical reasoning capacities comprised both of the ability and of the tendency to carry out key empirical processes in gathering and processing information, including formulating and asking empirical questions, generating and testing hypotheses, making inferences from data, and communicating those inferences to others. Critically, engaging with the world in this way requires both carrying out these empirical processes oneself, as well as evaluating whether others have done so effectively. These empirical habits of mind form the foundation not only of scientific literacy, but of critical engagement across many aspects of modern life and thus, as a society, we must work to ensure that our children develop these habits early in life (see Butler, 2020 for a comprehensive review).

Most of the research on children's understanding of evidence, per se, has focused on children's sensitivity to and inferences from various patterns of statistical evidence (see Gopnik, 2012; Schulz, 2012). However, children encounter a great deal of evidence not in the form of clear patterns of cause and effect, but in the form of verbal testimony—of others making factual claims about the world which children have to either trust or reject (Harris, 2012). Learning from others' testimony is inherently epistemologically challenging. Children must often infer whether information is reliable without actually *knowing* whether it is true.

Previous research has demonstrated that children's trust is selective rather than indiscriminate (see Harris, Koenig, Corriveau, &

* Corresponding author at: Department of Human Development and Quantitative Methodology, University of Maryland, 3942 Campus Drive, College Park, MD 20742, United States.
E-mail address: lpbutler@umd.edu (L.P. Butler).

https://doi.org/10.1016/j.cogdev.2020.100871
Received 20 March 2019; Received in revised form 26 February 2020; Accepted 27 February 2020

Jaswal, 2018), and children are judiciously skeptical of the informants they encounter (Mills, 2013). Recent research has also demonstrated that children are remarkably nuanced in the reliability judgments that they make about others, showing an early sensitivity to epistemic cues, such as others' accuracy and inaccuracy (e.g., Birch, Vauthier, & Bloom, 2008; Corriveau, Meints, & Harris, 2009), other's expertise (e.g., Danovitch & Keil, 2004, 2007) and possession of or access to the right kind of knowledge (Brosseau-Liard & Birch, 2011; Koenig et al., 2015; Kushnir, Wellman, & Gelman, 2008; Kushnir, Vredenburgh, & Schneider, 2013; Nurmsoo & Robinson, 2009; Pillow, 1989; Robinson, Butterfill, & Nurmsoo, 2011). Children also show a marked sensitivity to many socially relevant aspects of their informants, including authority (e.g., Michener & Burt, 1975) familiarity (e.g., Corriveau & Harris, 2009), whether the informant belongs to the same social group as them, be it one based on race, accent, gender, or even a minimal group manipulation (Chen, Corriveau, & Harris, 2013; Corriveau, Kinzler, & Harris, 2013; Hetherington, Hendrickson, & Koenig, 2014; Jaswal & Neely, 2006; Shutts, Banaji, & Spelke, 2010), dominance (Bernard et al., 2016), prestige (Chudek, Heller, Birch, & Henrich, 2012), and even attractiveness (Bascandziev & Harris, 2016). They also continually adjust their appraisals of informant reliability as they accumulate more information about an individual's accuracy (Ronfard & Lane, 2018). Further, they use this assessment to decide who to ask for different kinds of information (Danovitch & Keil, 2004, 2007; Kushnir et al., 2013; Mills, Legare, Grant, & Landrum, 2011), and to fine-tune how *much* weight to place on testimony from different individuals (Baum, Danovitch, & Keil, 2008; Corriveau & Kurkul, 2014; Mills & Keil, 2005, 2008). Finally, in situations where they are not evaluating others' testimony but rather crafting it themselves, they tailor their arguments to others' knowledge or lack thereof (Köymen, Rosenbaum, & Tomasello, 2014; Köymen, Mammen, & Tomasello, 2016).

This literature makes clear that young children are selective social learners, but it does not directly address a fundamental question about children's empirical understanding. All empirical approaches to knowledge rest upon the foundational principle that any factual claim must be based on sufficient knowledge and evidence, and that one ought to be skeptical of or reject claims that are not. We cannot always assess whether an individual claim is true, but we often can assess the validity of the *process* by which someone has acquired (or failed to acquire) the necessary knowledge to make a claim. Relatedly, an informant's prior accuracy or reliability, though broadly relevant to whether we trust what they tell us, is irrelevant if they do not have access to the right kind of information in the current situation. Imagine a scientist, presumably reliable within her domains of expertise, who makes a claim about an entirely different field of study. Rather than simply trusting the scientist as a universally knowledgeable informant, it is important to evaluate the empirical process by which she has come to that claim and evaluate it on the basis of whether it is supported by sufficient evidence.

Thus, in learning from others, children need to integrate their ability to judge reliable informants with their understanding of what constitutes sufficient verification for particular factual claims. Given the literature reviewed above, young children clearly possess the ability to do the former. Recent evidence also suggests that young children can evaluate whether an individual claim is acceptable when based on little evidence (Butler, Schmidt, Tavassolie, & Gibbs, 2018). But integrating these two capacities—making attributions about informants' reliability on one hand and evaluating whether informants' claims are based on sufficient evidence on the other—may be more challenging, especially when they may lead to conflicting judgments.

Very little research has targeted this integration question. Brosseau-Liard and Birch (2011) found that preschool-aged children could successfully use situation-specific cues (e.g., verifying the contents of boxes prior to making claims about them) and person-specific cues (e.g., history of accuracy in a labeling task) when deciding from whom they wanted to learn about either the identity or the name for hidden objects. However, we also need to ask *how* children weigh competing, epistemically-relevant cues to knowledge, when the role of person-specific cues is not pertinent, or only peripherally so. Two findings are relevant to this issue. First, Terrier, Bernard, Mercier, and Clément (2016) found that 3- and 4-year-olds, who generally favor informants who share their own gender, can override this and base their judgments on who has visual access. Second, Stephens and Koenig (2015) found that when an individual had established a previous history of reliability about semantic information (the name of novel objects), 3- and 4-year-old children were more likely to prefer that informant on later test trials than an informant who has established a previous history of reliability about episodic information (the location of a hidden object). Interestingly, this pattern held regardless of whether the test trials involved semantic or episodic testimony, suggesting that children may struggle to inhibit their reliance on prior accuracy when the cues to that accuracy are no longer relevant. However, neither of these papers provided a contrast between prior reliability and current access to knowledge.

2. The current research

In order to further investigate whether and at what age children understand the foundational empirical principle that any claim must be based on sufficient evidence, even if it comes from a previously reliable, trustworthy informant, we begin with an extremely simple scenario: one in which one individual has access to and generates the right kind of evidence to make a claim while another, in contrast, makes a claim without having generated that evidence. The current research makes several steps forward in this line of inquiry. First, in prior work (e.g., Brosseau-Liard & Birch, 2011), children chose between an epistemically-relevant cue (looking in a box) and an epistemically-irrelevant cue (standing on a box). The current experiments provided a stricter test, forcing children to select an informant on the basis of two epistemically-relevant cues (looking *at* vs. looking *in*). This is an important distinction. Young children may understand that visual access is necessary, and preferable to a complete lack of visual access, but they may have a relatively blunt understanding of what constitutes sufficient in each given case. Indeed, one might plausibly be able to make an educated guess simply by looking at the outside of a container, depending on the situation. Thus, the distinction we use (looking *at* vs. looking *in*) presents a relatively conservative test of children's understanding of what counts as sufficient verification. Prior work manipulating visual access has either contrasted epistemically-relevant vs. irrelevant actions (Brosseau-Liard & Birch, 2011), or provided children with explicit verbal scaffolding of who did and did not have visual access (Terrier et al., 2016). It is entirely conceivable that children may show age differences in whether or not they can make the more fine-grained distinction between sufficient and insufficient epistemically-relevant actions. Using this more conservative manipulation of visual access may thus provide a more accurate picture of the development of children's understanding of verification.

Second, the current research provides a more objective test of children's understanding. Most, if not all, of the research on children's selective social learning focused on their conclusions about what is true, on the basis of different epistemic and social cues. That is, children are tasked with learning, and experiments test whether they selectively learn from a reliable or an unreliable source. However, children's own learning is not the same as their understanding of how one *ought* to learn. Findings from several literatures are relevant here. First, children between 3 and 8 appear to be constructing an understanding of what learning is and how it works, and this work has emphasized the importance of distinguishing between children's reasoning about how they themselves learn, how others learn, and how teaching works (Sobel & Letourneau, 2015, 2016, 2018) Second, work on children's epistemological reasoning (e.g., Kuhn, Cheney, & Weinstock, 2000) suggests that younger children begin as objectivists about knowledge—that is they believe there are objective truths in every case—and that assertions are direct reflections of the world, even in subjective cases like personal taste or preference. Later in development, this yields to subjectivity—believing in the primacy of opinions even in cases with an objective truth Finally, the objectivist and subjectivist perspectives are integrated. Critically, by around age 4, children begin to understand that assertions or claims are reflections of individuals' beliefs, and those beliefs may be more or less valid with respect to reality (see Kuhn, 2000). Thus, children at this age can reflect on the relation between knowledge and claims. Finally, work in the area of normative reasoning suggests that around age 3, children develop the ability to reason about how one *ought* to behave, and use normative language when others engage in non-normative behavior (Tomasello & Vaish, 2013; Tomasello, 2018). Though focused on different questions than the current research, that work is relevant to children's thinking about how knowledge *ought* to be gathered and conveyed. Indeed, by age 3, children can understand the normative dimensions of false claims (Fedra & Schmidt, 2019) In order to more directly ground our work in the question of children's understanding of how one *ought* to gather information, we asked if children could use their understanding of verification on behalf of a third party, that is to help guide their choice of whether a naïve third-party ought to learn from an informant making verified or unverified claims.

Third, the current research asked not only whether children could use their understanding to guide their selection of whom a third-party individual ought to learn from, but their explicit understanding of *why* the third-party individual ought to learn from one informant and not the other. To assess this explicit understanding, we asked children to justify their forced-choice responses, allowing us to examine the explicit reasoning underlying their choices. Most of the literature on selective trust (see Harris et al., 2018) has relied solely on forced-choice measures, which do not directly assess children's explicit understanding of why one source out to be trusted over another. However, the work on argumentation and joint decision-making suggests that, at least by age 5, children are able to provide explicit, valid justifications in their reasoning and decision-making (Köymen et al., 2014, 2016; see also Mercier, 2011 for a broad review of children's argumentation skills). Particularly relevant to the current work, Köymen et al. (2014, 2016) found that children flexibly make use of who knows what in order to guide their arguments, suggesting an understanding that successful claims are grounded in knowledge or the lack thereof. Given this, adding a measure of children's explicit reasoning may shed important light on the developing understanding of verification.

Finally, the current work investigates a broader developmental range, from age 3 to age 7. In conjunction with the factors outlined above, especially the inclusion of the explicit reasoning prompt, this provides a more comprehensive picture of possible age differences than previously demonstrated in the literature.

In the current experiments, children were shown puppet shows similar to those depicted in the Brosseau-Liard and Birch (2011) studies. In Experiment 1, 4- to 7-year-olds had to select between an informant who had performed the right epistemically relevant action (looking in a box before making a claim about what was inside it) an epistemically relevant but insufficient action (looking at the box and making a claim about what was inside it), and then had to justify those responses. In Experiment 2, we adapted this paradigm to assess 3-year-olds' understanding of this distinction. Finally, in Experiment 3, 4- to 7-year-olds saw the same puppet shows as in Experiment 1, but in this case, the two informants had previously established a history of accuracy or inaccuracy by providing correct or incorrect labels for familiar objects. On half of the test trials, the puppets' epistemically relevant actions were congruent with their past reliability (the previously accurate puppet also verified her claim, and the previously inaccurate puppet failed to do so) or incongruent with past reliability (the previously accurate puppet failed to verify her claim, while the previously inaccurate puppet did so successfully).

3. Experiment 1

Experiment 1 served to replicate prior work (e.g., Brosseau-Liard & Birch, 2011), as well as to extend those investigations by more tightly controlling the epistemically relevant behaviors of the two informants, testing this in an objective, third-party context, and assessing children's explicit reasoning. In this study, 4- to 7-year-old children were shown a digitalized puppet show depicting two puppets, one who consistently made claims only after verifying their accuracy, and one who consistently failed to do so. After each trial, children were asked to indicate which of the two informants did a better job teaching a third, naïve puppet, and to provide justifications for their selections. Thus, we examined whether 4- to 7-year-old children could successfully arbitrate between informants who either did or did not verify evidence prior to making claims, and whether they would use that arbitration to select from which of two informants a naïve, third-party learner should learn.

We had three main predictions: First, based on previous research, we predicted that children at all ages would choose the puppet who verified their claims over the puppet who failed to verify their claims at significantly more often than chance, with increasing success across ages. Second, we predicted that the extent to which children's justifications for their selections explicitly referenced the process of checking or verifying would increase with age and would correlate significantly with their selections on the forced choice test question.

3.1. Participants

In Experiment 1 we tested 24 4- and 5-year-old children (Mean age = 62 months; 12 males, 12 females) and 24 6- and 7-year-old children (Mean age = 81 months; 12 males, 12 females). Children were recruited from a local children's museum in a medium-sized mid-Atlantic city. Although the sample was largely upper-middle class, children represented a range of ethnic backgrounds (50 % Caucasian/White, 25 % Biracial/Mixed, 11 % African American/Black, 8% Asian/Pacific Islander, 6% Hispanic/Latino/a). Some participants opted not to provide ethnic/racial information.

3.2. Procedure

Children were shown a digitally rendered puppet show (presented as PowerPoint slides on a laptop). Children were first introduced to a naïve puppet ("Bear") and were told that "Bear is from a country really far away and doesn't know much about the things we have here or the things we do here," and were asked whether they could help Bear learn the right things. Then children were introduced to two more puppets ("Mouse" and "Sheep") and were told that they would teach Bear some things. However, children were told, "sometimes they'll do a *good* job teaching Bear, and tell him the *right* things, and sometimes they *won't* do such a good job teaching Bear and tell him the *wrong* things." After each trial, children were asked to indicate which of the two informants they believed Bear should listen to, and to provide justification for their selection.

3.2.1. Accuracy trials

Children saw four accuracy trials as a familiarization for the test phase, during which they watched videos of Mouse and Sheep accurately or inaccurately labeling familiar objects (e.g., in reference to a soccer ball, "That's a ball!" vs. "That's a cup!"). After each trial, children were asked to choose which of the two informants Bear should listen to (e.g., "Who should Bear listen to, Mouse or Sheep?"), and to provide justifications for their selections (e.g., "Alright, why should Bear listen to [child's choice]?"). Both the order of trials and the identity of the accurate and inaccurate puppets were counterbalanced across participants.

3.2.2. Verification trials

Following the accuracy trials, children were then introduced to two *new* puppets ("Pig" and "Duck"). Again, children were told that these two informants would teach Bear some things, and that sometimes they would do so reliably and sometimes unreliably. On the four verification trials, children saw videos of Pig and Duck making conflicting claims about the contents of boxes. On each trial, one informant looked *inside* the box (verified claim), while the other looked *at* the box but did not open it (unverified claim). After each trial, children were again asked to select which of the two informants Bear should listen to, and to provide justifications for their selections. Both the order of trials and the identity of the puppet verifying or failing to verify were counterbalanced across participants.

3.3. Coding

Children's performance on the accuracy and verification trials were coded as either correct (selecting the puppet who used the correct label on the accuracy trials; selecting the puppet who looked inside the box on the verification trials) or as incorrect (selecting the puppet who used the incorrect label on the accuracy trials; selecting the puppet who failed to look inside the box on the verification trials). Children thus received a total score of 0–4 for each type of trial.

We also coded children's explanations for their choice on the verification trials (their justifications for their forced-choice selections) to assess the sophistication of the reasoning underlying their selections. To do so, children were given a score from 0-3. Children received a 3 if they explicitly referenced verification (e.g., "because he looked in the box," "because he opened it," "because he saw what's in there"). Children received a 2 if they failed to explicitly reference verification per se, but explicitly referenced knowledge (e.g., "because he knows what's in the box"). Children received a 1 if they made only an implicit reference to the puppet's knowledge, accuracy, or confident assertion (e.g., "because he said the right thing," "because he said it was shoes," "because he's right"). Children received a 0 for all other responses.

3.4. Results

3.4.1. Accuracy trials

Although the accuracy trials served primarily as a familiarization in this study, we nevertheless analyzed whether children's success at choosing the accurate informant varied with age. A repeated-measures ANOVA, with age group as a between-subjects factor, trial as a within-subjects factor, and forced-choice selection as a dependent measure, revealed a non-significant effect of age group ($F(1,46) = 2.76$, $p = .103$, partial $\eta^2 = .057$). Consistent with prior research, children performed above chance on the accuracy trials ($M = 3.88$ out of 4 trials, $SD = .531$, $t(47) = 24.47$, $p < .001$, $d = 3.54$).

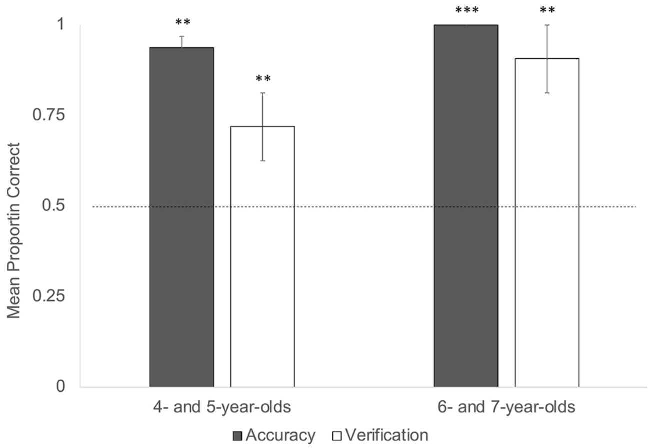

Fig. 1. Mean proportion of accuracy and verification trials on which children in each age group selected the correct informant in Experiment 1. Error bars represent +/- 1 SEM. The dashed line represents chance responding (50 %). ** $p < .01$; *** $p < .001$.

3.4.2. Verification trials

Our key measure of interest for this study was children's success on the verification trials, and its relation to their explicit reasoning (see Fig. 1). In order to assess this, we conducted a repeated-measures ANOVA, with age group as a between-subjects factor, trial as a within-subjects factor, mean explanation score as a covariate, and forced-choice selection as a dependent measure. This analysis revealed a non-significant effect of age group ($F(1,45) = 2.17$, $p = .148$, partial $\eta^2 = .046$), and a significant effect of their mean explanation score ($F(1,45) = 10.4$, $p = .002$, partial $\eta^2 = .188$). Both 6- and 7-year-olds ($M = 3.63$, $SD = .647$, $t(23) = 12.31$, $p < .001$, $d = 2.51$), and 4- and 5-year-olds ($M = 2.88$, $SD = 1.45$, $t(23) = 2.95$, $p = .007$, $d = .607$) selected the puppet who had verified their claims above levels expected by chance. This suggests that they understood that the naïve puppet ought to learn from the informant who had verified their claims by checking the contents of the containers, and that this understanding is related to their explicit understanding of the process of making knowledge-based claims.

Furthermore, results showed some evidence of age differences in the consistency of children's responses across the forced choice trials. Thirteen out of 24 (54.2 %) of 4- and 5-year-olds chose correctly on all 4 trials, and 15 out of 24 (62.5 %) chose correctly on at least 3 trials. Seventeen out of 24 (70.1 %) of 6- and 7-year-olds chose correctly on all 4 trials, and 22 out of 24 (91.7 %) chose correctly on at least 3 trials.

In order to further assess children's explicit understanding of verification, as measured by the quality of their explanations, and its relation both to age and to their performance on the force-choice task, we conducted several additional analyses. First, to assess the relation between age and children's explanations, we ran an additional repeated measures ANOVA, with age group as a between-subjects predictor and explanation score as a dependent measure. This analysis revealed a significant effect of age group ($F(1,45) = 4.42$, $p = .041$, partial $\eta^2 = .089$). Six- and 7-year-olds had higher average explanation scores ($M_{6\text{-}7\text{-year-olds}} = 1.37$ out of 3, $SD = 1.11$) than the 4- and 5-year-olds ($M_{4\text{-}5\text{-year-olds}} = .771$ out of 3, $SD = .847$, $t(46) = 2.09$, $p = .042$, $d = .603$).

Finally, we assessed the alignment between children's explanations and their forced choice responses. To do this, we coded for each trial whether their forced-choice and explanations matched. A trial was coded as a match either if they chose correctly and gave an explanation that was coded as a 2 or 3 (explicitly referencing either verification or knowledge) or if they chose incorrectly and gave an explanation that was coded as a 0 or 1 (only implicitly referencing something relevant about the puppet's mental states or accuracy, or providing an irrelevant response). A trial was coded as a mismatch either if children answered correctly but gave an explanation that was coded as a 0 or 1, or if they chose incorrectly but gave an explanation that was coded as a 2 or 3. Overall, children matched on 1.83 out of 4 trials, and mismatched on 2.17 out of 3 trials. However, no children in the entire sample chose incorrectly on the forced-choice while providing a higher-level explanation—all the mismatches were ones in which children chose correctly but did not provide a higher-level explanation. This suggests that children's ability to explain their reasoning by explicitly referencing the relevant mental states of the puppets was sufficient, but not necessary, for them to choose the correct puppet on the forced-choice task. This is consistent with the conclusion that children's implicit understanding of verification may precede their ability to articulate it.

3.5. Discussion

The results of Experiment 1 demonstrate that by age 4, children are at least somewhat capable of understanding that epistemic claims need to be based on evidence and can extend this understanding beyond their own learning aims to guide a naïve, third party towards choosing a reliable source. However, there were clear age-related differences both in terms of how reliably children indicated that the third party should listen to the individual who had verified their claims, as well as to what extent they were able to explicitly justify those selections in terms of verification behavior. Further, the extent to which they used verification to judge who a third party should learn from was related to the quality of their explanations.

Critically, in the current research, children had to make these selections without knowing the truth value of any of the individuals' statements. Children never saw what was inside the containers, and thus they could only base their judgments on the evidentiary process by which individuals arrived at their claims. Children were able to do so by age 4, but show development in the sophistication of this skill between 4 and 7. In Experiment 2, we test whether children as young as 3 show evidence of this understanding.

4. Experiment 2

4.1. Participants

In Experiment 2 we tested an additional 24 3-year-old children, (Mean age = 42 months; 12 boys, 12 girls). As in Experiment 1, children were recruited from a local children's museum in a medium-sized mid-Atlantic city. Although the sample was largely upper-middle class, children represented a range of ethnic backgrounds (58 % Caucasian/White, 25 % Biracial/Mixed, 13 % African American/Black, and 4% Asian/Pacific Islander).

4.2. Procedure

The procedure was identical to Experiment 1, with one modification. Pilot testing indicated that these younger children struggled to focus through the entire task, and thus children in Experiment 2 saw an abbreviated version, with only 3 accuracy and 3 verification trials.

4.3. Results

4.3.1. Accuracy trials

As with older children, 3-year-olds performed above chance on the accuracy trials, selecting the accurate over the inaccurate informant (M = 1.92 out of 3, SD = .929, t(23) = 2.20, p = .038, d = .452).

4.3.2. Verification trials

As in Experiment 1, the key measure of interest for this study was children's success on the verification trials, and its relation to their explicit reasoning. In order to assess this, we conducted a repeated-measures ANOVA, with trial as a within-subjects factor, mean explanation score as a covariate, and forced-choice selection as a dependent measure. This analysis revealed only a significant effect of children's mean explanation score (F(1,22) = 9.83, p = .005, partial η^2 = .309). Overall, however, 3-year-olds selected the verifying puppet only marginally more often than chance (M = 1.86 out of 3, SD = .899 t(23) = 2.04, p = .053, d = .400). In order to assess potential developmental differences within this age group, we ran several bivariate correlations between age, in months, and both the number of verification trials correct and the number of trials referencing verification. There was a marginal correlation between age and number of verification trials correct (r = .376, p = .070), and a significant correlation between age and their mean explanation score (r = .426, p = .038).

Further, children's responses across the forced-choice trials were not as consistent as those of older children. Only 6 out of 24 children (25 %) answered correctly on all 3 trials, and 17 out of 24 (70.1 %) chose correctly on at least 2/3 of trials.

Finally, children showed a similar level of alignment between their explanations and their forced-choice responses as seen in older children. Overall, 3-year-olds matched (as described above) on 1.13 out of 3 trials and mismatched on 1.87 out of 3 trials. Of those who mismatched, children never chose the incorrect puppet on the forced-choice question while providing a higher-level explanation. Thus, as with older children, their ability to articulate their reasoning about verification seems to be sufficient, but not necessary, for them to select the verifying puppet, suggesting that their explicit understanding may be preceded by an implicit understanding.

4.4. Discussion

These results suggest that 3-year-olds are just beginning to understand the importance of verification in making claims, at least as assessed by the current measures. However, as in the older children, the quality of their explanations both correlated with age, as well as with their forced-choice performance.

Taken together, the results of Experiment 1 and 2 show substantial development in children's understanding of verification as a necessary condition for making factual claims, both as assessed by their forced-choice selections on behalf of a naïve third party, as well as by their developing ability to provide explicit justifications of those choices.

However, children hardly ever receive information in a vacuum. It is conveyed by individuals, with a variety of both social and epistemic motives. In many cases, it is likely adaptive to judge factual claims on the basis of their source—that is, information from a trustworthy source is taken as true, even without any information about its legitimacy. By and large, this form of trust is likely to lead to the uptake of accurate information. However, this social learning mechanism presents a danger, in that generally reliable sources, though trustworthy overall, cannot be assumed to be reliable at all times, in all domains, without constraints. In particular, in cases where a previously reliable individual makes a plausible, potentially true claim about something they in fact lack sufficient knowledge of, we need to be able to inhibit our general trusting stance *in general* in favor of skepticism *in the moment*. Children's ability to integrate their reasoning about informant reliability with their understanding of what constitutes sufficient evidence for particular factual claims is addressed in Experiment 3.

5. Experiment 3

Experiment 3 builds on the first two experiments by asking an additional critical question. Specifically, as they develop an understanding that factual claims must be based on access to sufficient knowledge, are children able to use that understanding to evaluate such claims even when access to knowledge conflicts with the previous reliability of a given individual? As mentioned in the introduction, we cannot simply trust, say, a scientist as a universally knowledgeable informant about all things, even in her subject area, but rather must evaluate the process by which she has weighed evidence and made claims. Consider several further examples from modern life. A lawyer is, by necessity, highly educated and assumed to be an expert in legal issues. However, we do not rely on lawyer's general expertise and presumed trustworthiness in making decisions about guilt and culpability. Rather, we rely on an impartial assessment based on the "facts of the case," and we rely on individuals who have a legitimate claim to knowledge (i.e., eyewitnesses) to provide compelling evidence. With all its pitfalls, the legal system is predicated on the assumption that specific knowledge and evidence outweigh general trustworthiness as a matter of principle.

On a more everyday level, consider two mechanics trying to diagnose a problem with your car. You take it to your trusted mechanic, on whom you have relied for many years, but they are busy, do not take the time to look carefully at your vehicle, and give you a guess about what they think the problem is. You decide to get a second opinion, at a new mechanic about whom you have not heard great things. They spend half a day inspecting all the systems in your vehicle and arrive at a completely different conclusion. Who do you trust? Even though you might feel badly about it, chances are your previous mechanic just lost your business, at least for this repair, and with good reason. Despite their previous reliability, they were negligent in this case, while the new mechanic did their due diligence. You are better off going against your initial impulse, and trusting the new mechanic, at least in this case.

To return to children's reasoning, the scenario children encountered in Experiments 1 and 2 is broadly similar to the mechanic example, but with no prior history of reliability, making the decision rather straightforward. The question asked in Experiment 3 more closely parallels the mechanic example, asking how children grapple with the dilemma of trusting a previously reliable source who is being negligent in the moment, versus a previously unreliable source who has the right kind of knowledge for the current situation, children are able to choose basis on current access to knowledge.

To assess this question, the procedure in Experiment 3 was identical to that of the previous experiments, except that the informants making verified or unverified claims were the same ones who had previously labeled familiar objects accurately or inaccurately, thus giving them a baseline level of reliability against which their current verification (or failure to verify) could be measured. Critically, as in Brosseau-Liard and Birch (2011), on some trials, the puppets' verification or lack of verification was congruent with their previous accuracy (i.e., the accurate puppet verified her claims and the inaccurate puppet failed to), and on other trials, the puppets' verification or lack of verification was incongruent with their previous accuracy (i.e., the inaccurate puppet verified her claims and the accurate puppet failed to). These incongruent trials are particularly revealing, as they directly ask the question of whether children can inhibit their tendency to base trust decisions on prior reliability when an individual's current negligence ought to outweigh it.

The current work closely parallels that by Brosseau-Liard and Birch (2011) but adds to it in several important ways, as detailed in the introduction. First, the distinction between the behaviors of the two informants was subtler. Both informants performed an epistemically relevant behavior, but only one did enough to gain the right kind of information (looking *in* as opposed to looking *at* the box). Second, we asked children to make their selections on behalf of a naïve third party, providing a measure of their objective understanding of the importance of verification. Third, we asked children to explicitly explain their forced choices, providing a measure of their explicit understanding of the relation between verification and claims.

We had two key predictions: First, we predicted that children would be less likely to successfully select the puppet who had verified their claims, when they had previously labeled familiar objects inaccurately. Second, we predicted that this would show age-related differences. Specifically, we predicted an age by condition interaction, with older children more likely to consistently select the verifying puppet specifically on incongruent trials, in which the previously accurate informant failed to verify her claims, and the previously inaccurate informant did so successfully. Additionally, we predicted that, as in Experiment 1, children's ability to explicitly justify their selections would increase with age and would be correlated with their success at selecting the verifying informant.

5.1. Participants

In Experiment 3, we recruited an additional 32 4- and 5-year-old children (Mean age = 59.3 months; 16 males, 16 females) and an additional 32 6- and 7-year-old children (Mean age = 83.4 months; 16 males, 16 females). Children were recruited from the same local children's museum, located near a medium-sized mid-Atlantic city. This sample was similarly upper-middle class, and also represented individuals from a range of ethnic backgrounds (53 % Caucasian/White, 20 % Biracial/mixed, 12 % African American/Black, 11 % Asian/Pacific Islander, and 4% Hispanic/Latino/a).

5.2. Procedure

The procedure was identical to that of Experiment 1, with one key change. Instead of the puppets in the accuracy trials being different puppets from those in the verification trials, this time they were the same. That is, Pig and Duck engaged in accurate and

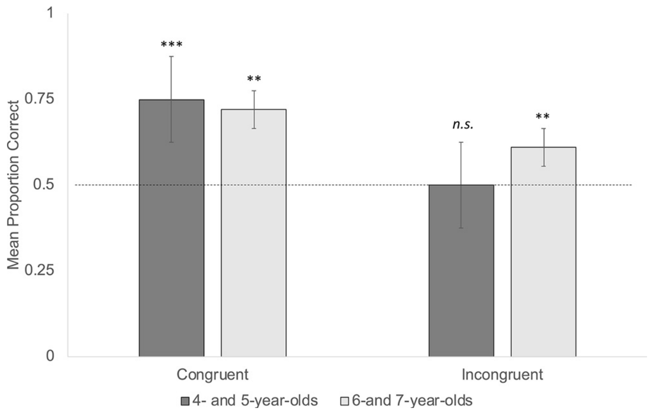

Fig. 2. Mean proportion of verification trials on which children in each age group selected the verifying informant across the congruent and incongruent trials in Experiment 2. Error bars represent +/- 1 SEM. The dashed line represent chance responding (50 %). ** $p < .01$; *** $p < .001$.

inaccurate labeling, and then Pig and Duck verified or failed to verify their claims about the contents of novel containers. Thus, before the verification trials, these puppets had established a history of reliability, with one having accurately labeled the familiar objects, and one having inaccurately labeled them. Two of the four verification trials were congruent (the previously accurate informant also verified, while the previously inaccurate informant failed to), and two trials were incongruent (the previously accurate informant failed to verify, and the previously inaccurate informant verified). The order of trials and trial types, as well as the identity of the previously accurate or inaccurate puppet were counterbalanced across participants.

5.3. Results

5.3.1. Accuracy Trials.

As in Experiment 1, children at both ages chose the accurate puppet over the inaccurate puppet on accuracy trials significantly above chance ($M_{4\text{-}5\text{-year-olds}} = 3.38$, $SD = 1.21$, $t(31) = 6.45$, $p < .001$; $M_{6\text{-}7\text{-year-olds}} = 3.97$, SD = .177, $t(31) = 63.0$, $p < .001$).

5.3.2. Verification trials

The key measure of interest for this study was children's success on each of the two types of verification trials (congruent vs. incongruent), and its relation to their explicit reasoning (see Fig. 2). In order to assess this, we conducted a repeated-measures ANOVA, with age group as a between-subjects factor, condition and trial as within-subjects factors, mean explanation score as a covariate, and forced-choice selection as a dependent measure. This analysis revealed a significant interaction between age group and condition ($F(1,61) = 5.32$, $p = .024$, partial $\eta^2 = .080$, and a significant effect of their mean explanation score ($F(1,61) = 12.0$, $p = .001$, partial $\eta^2 = .164$).

The 4- and 5-year-olds performed above chance on congruent trials ($M = 1.56$, $SD = .716$ $t(31) = 4.45$, $p < .001$, $d = .782$), but showed at chance performance on incongruent trials ($M = 1.03$, $SD = .822$, $t(31) = .226$, $p = .823$, $d = .037$). They thus performed significantly better on congruent than incongruent trials, ($t(31) = 3.16$ $p = .004$, $d = .688$). Six- and 7-year-olds performed above chance on both congruent trials ($M= 1.44$, $SD = .716$, $t(31) = 3.26$, $p = .003$, $d = .615$) and incongruent trials ($M= 1.38$, SD = .609, $t(31) = 3.48$, $p = .002$, $d = .624$).

Furthermore, the consistency of children's performance across the trials varied by trial type. For congruent trials, 24 out of 32 (75 %) of 4- and 5-year-olds and 19/32 (59.4 %) of 6- and 7-year-olds answered both trials correctly. Thirty out of 32 (93.8 %) of 4- and 5-year-olds and 27/32 (84.4 %) of 6- and 7-year-olds answered at least 1 trial correctly.

For incongruent trials, performance was more mixed. Eleven out of 32 (34.4 %) of 4- and 5-year-olds and 14 out of 32 (43.8 %) of 6-7-year-olds answered correctly on both trials. Twenty-two out of 32 (68.8 %) of 4- and 5-year-olds and 29 out of 32 (90.6 %) of 6- and 7-year-olds answered at least 1 trial correctly.

In order to further investigate the development of children's explicit reasoning about verification, we conducted a repeated measures ANOVA, with age group as a between subject factor, condition and trial as within-subject factors, and children's explanation score as a dependent measure. The only significant effect revealed was a trial by age group interaction ($F(1,62) = 6.13$, $p = .016$, partial $\eta^2 = .090$). The 4- and 5-year-olds showed no difference in their mean explanation scores between the first ($M = .672$, $SD = .819$) and second ($M = .703$, $SD = .932$) of each trial type ($t(31) = .263$, $p = .794$, $d = .035$), but 6- and 7-year-olds showed significant improvement from the first ($M = .797$, $SD = 1.05$) to the second of each trial type ($M = 1.39$, $SD = 1.37$, $t(31) = 3.07$, $p = .004$, $d = .486$). This suggests that as children saw the contrast between the first two exemplars of the two trial types, they were better able to articulate their explanations for their forced choices.

Finally, as in Experiments 1 and 2, we assessed the consistency of children's forced-choice responses and their explicit explanations. Overall, as in the previous studies, nearly children at both ages either matched in their forced-choice success and the level of the explanation ($M = 2.08$ out of 3, $SD = 1.34$), or made correct selections on the forced-choice questions but failed to give a higher-level explanation (M = 1.84 out of 3, SD = 1.35). Only two children ever gave a higher-level explanation without also choosing correctly. Thus, as in Experiments 1 and 2, children's implicit understanding as referenced by their forced-choice selections may precede their explicit understanding.

5.4. Discussion

The results of Experiment 3 illustrate clear age-related differences in children's understanding of verification. Children in Experiment 1 were able to use verification behavior to guide their choice about who a naïve individual should learn from as early as age 4, and this effect replicated in the current study on congruent trials, when the informants' verification behaviors were consistent with their prior accuracy. But on incongruent trials, when informants' prior accuracy and current verification behavior were in conflict, children younger than age 6 did not reliably select the verifying puppet. This shows that children are working to integrate their reasoning about the evidentiary process by which a claim is made with their reasoning about other relevant cues to reliability, and this integration process appears to be nuanced and challenging. Indeed, past reliability is probably a relatively good heuristic for deciding who to trust, and previous work (e.g., Brosseau-Liard & Birch, 2011; Nurmsoo & Robinson, 2009) has indicated that children may actually weigh this cue relatively equally to other critical epistemic cues. And yet, as we have described above, there are cases in which one must set aside one's knowledge of who an informant is and whether they are generally reliable in favor of their current epistemically-relevant behavior. The most expert person in the world is not omniscient and cannot base a claim on knowledge they do not have.

6. General discussion

Much of what we know about the world around us is based purely on what others tell us (Harris, 2012). But in depending on testimony from others, a crucial challenge is using epistemic reasoning to assess whether others' testimony about unknown domains of knowledge or sources of evidence is likely to be reliable. The critical thinking skills required to assess the validity of communicated information is essential to children both as they gain knowledge in various academic disciplines, and to the construction of a reliable conceptual understanding of the world more broadly. In order for children to become skeptical, critical consumers of information, they must be able evaluate the quality of the evidentiary actions underlying specific acts of testimony. Moreover, children need to integrate their understanding of the factors that indicate that an individual is a reliable source of information, with their understanding of what constitutes sufficient evidence for particular factual claims. They must also learn to privilege the gathering of sufficient evidence above and beyond, and sometimes regardless of, the general trustworthiness of the individual making a particular claim.

The current research suggests that, by age 4, the majority of children recognize that one ought to accept verified claims over unverified ones, and use this understanding to guide a naïve learner, Moreover, the consistency with which children use this understanding in selecting information increases with age, and is related to a developing ability to explicitly articulate why they made those selections. Further, although younger children were not able to correctly select information on the basis of verification at above-chance levels, their ability to do so was similarly related to their ability to justify their choices. Finally, there were indications of developmental differences in whether children were able to successfully integrate their reasoning about informant reliability with their evaluation of the evidence underlying particular claims. When faced with a conflict between individuals' past reliability and their current faithfulness or negligence in verifying information, only older children were above chance in choosing the informant who gathered sufficient information, even if they had previously in inaccurate. Four-and 5-year-olds were able to use verification to evaluate from whom a naïve third-party should learn when an informants' verification behavior was consistent with their prior accuracy, thus presenting clear-cut, non-conflicting cues to reliability. When their verification conflicted with their prior accuracy, however, children in this age group performed at chance. In contrast, 6- and 7-year-olds' use of verification was more robust, and they were reliably able to choose the correct, verifying informant above chance even when that informant had a previous history of inaccuracy. This is consistent with the findings from Brosseau-Liard and Birch (2011), who found that when informants made verbal claims about the contents of the containers, they were at chance at selecting the previously accurate puppet who failed to look in the box or the previously inaccurate puppet who did look in the box.

The current research makes several contributions to the nascent literature on children's ability to integrate their evaluation of informant reliability and their evaluation whether claims are based on sufficient verification, when tasked with selective social learning. First, as mentioned previously, this is one of the first investigations that directly targeted this integration issue, asking whether children understand that all factual claims must be based on sufficient evidence, even if they are made by a previously reliable individual. Second, the inclusion of a broader age range than in previous studies gives us a more comprehensive developmental picture than previous studies have been able to provide and illustrates that the integration of these two capacities is challenging for young children. Third, the addition of open-ended explanations means that these studies provide not only a window not into children's *implicit* understanding, as measured by the forced-choice selection measure, but into their *explicit* reasoning about why one ought to select one informant over the other. As the results of all three studies show, the quality of children's explanations was a consistent predictor of children's forced-choice answers. Although children sometimes chose correctly without being able to explain why, the reverse was never true, suggesting that their implicit understanding may precede their ability to provide explicit verbal justification of their reasoning.

Nevertheless, there remain important open questions that must be addressed in future research. The first concerns the developmental mechanisms underlying the results of these studies. Although showing clear development in this age range, the current work does not directly address the questions of what precisely underlies this development. There are several possibilities. First, although children do reliably pass tests of false belief by age 4 (Wellman, Cross, & Watson, 2001), there is still considerable development in children's social cognitive skills over this developmental span (see Wellman & Liu, 2004). To the extent that reasoning about verification may rely on social cognitive inferences about informants' knowledge and belief states, children's ability to reason about evidence may be relative to their capacity to engage in these cognitive skills. Second, children's executive function skills—including working memory and cognitive control, among others—are also developing during this time period (Diamond, 2013; Zelazo et al., 2003), and thus to the extent to which these tasks require considering multiple perspectives about the same question, individual differences in executive function skills may help explain some of the variance in children's reasoning about verification, as well as variability in their selection of the verifying informant on behalf of a third party. This seems especially true for Experiment 3, in which children must consider both the prior history of accuracy and the current verification behavior, and must inhibit reasoning about accuracy, which they have considered for 4 trials prior to this test phase, in favor of reasoning only about verification.

Second, the current research, as well as that on which it builds, has focused primarily on the rudimentary, basic cognitive foundations for reasoning about the link between evidence, verification, and knowledge-based claims. Specifically, the current work focused on the fundamental principle underlying all empirical approaches to reasoning, namely that one must have evidence to make factual claims, and that one may only make claims for which one has sufficient evidence. However, more work remains to be done in order to clearly connect these basic building blocks to the broader issue of scientific reasoning and its development. There is a wealth of evidence for preschool-aged children's ability to use patterns of evidence in order to support inferential conclusions, most notably from the literature on causal inference (see Gopnik, 2012; Schulz, 2012, for in-depth reviews). However, the ability to make inferences on the basis of causal evidence, and the ability to *understand* the inferential process, are not one and the same. Much of the work on children's understanding of the connection between scientific evidence and inferential conclusions focuses on older children and is less sanguine about younger children's scientific reasoning capacities (see Kuhn & Dean, 2004; Kuhn, Katz, & Dean, 2004, for in-depth reviews). However, preschoolers do show some ability to reason in more explicitly scientific ways. For example, by sometime between the ages of 6 and 8, children are able to differentiate between notions that might be true (i.e., hypotheses), and actual evidence for those hypotheses (Sodian, Zaitchik, & Carey, 1991). By age 6, and potentially even as young as 4, children can correctly reason about what someone will conclude from a pattern of false evidence (Koerber, Sodian, Thoermer, & Nett, 2005; Ruffman, Perner, Olson, & Doherty, 1993). And between ages 3 and 6, children are able both to systematically test hypotheses and to provide evidence to help correct someone else's false conclusion (Köksal-Tuncer & Sodian, 2018). This work poses several important questions for further research integrating these literatures. First, can children use their understanding of both causal relations and the logic of verification to reason about who is producing the right kind of evidence to reach a particular inferential conclusion, and who is producing relevant but insufficient evidence? Can children distinguish between inferences that are conclusively supported by a pattern of evidence, from ones that are consistent with the data but about which there remain confounds? That is, can children recognize when individuals overstate their conclusions, and draw inferences that go beyond what they ought to conclude form a given set of data?

Further, it will be important to investigate how children integrate this understanding with their developing social cognitive skills in order to reason about why individuals might make unverified claims, and whether individuals are making claims for less prosocial motivations such as persuasion or deceit. Do they understand that scientific reasoning is predicated on the notion that one is transparent and truthful in producing and conveying evidence? That is, do they grasp that making unsubstantiated claims is not only incorrect in terms of the conclusions one can draw, but unethical and potentially harmful to others (cf., Fedra & Schmidt, 2018)? Claims, especially misleading or unverified ones, may be made for a variety of social motives, and thus future work should investigate how children reason about those motives.

Finally, more work is needed to move these lines of investigation beyond scientific reasoning to other areas of learning. Indeed, learning in any domain, be it science, history, or social studies, relies on claims made and lessons drawn from a body of gathered information. Although what it means to "verify" information varies greatly across subject matters, understanding the process of inquiry and knowledge-generation in different academic disciplines is critical to successful engagement in those disciplines. To put it another way, by age 4, children understand the bedrock principal that you cannot say what you do not know. But learning how to learn in different fields requires both understanding how one *knows* something and what counts as verification in that field, as well as fostering the empirical habit of mind of continually asking, "how do we know?" Given the basic foundational capacities that children display in the current research, future work should investigate how best to capitalize and build on those capacities as children begin formal education across different subject areas.

Children and adults alike are increasingly inundated with information from sources of varying degrees of reliability and trustworthiness. We rely heavily on what others tell us in order to learn about the world and assessing the veracity of various sources is an important life skill. This is especially true with respect to critical thinking and scientific literacy, where insufficiently verified claims (e.g., that vaccines cause autism) are often made, adopted, and spread, leading to misconceptions and misunderstanding about important issues. Indeed, fostering scientific literacy, and combatting its possible decline, has been termed a "shared responsibility" by the Smithsonian Institution (Clough, 2010).

The ability to engage critically with the world, to actively question what we see and are told, and to use that questioning to guide further inquiry, is thus of ever-increasing importance in modern society. As a step towards addressing this issue, the current research aimed to start building a coherent picture of how and when children develop the ability to critically evaluate empirical claims. Ultimately, future research has the potential to generate clear targets for programs and interventions that seek to foster the development of these skills at the earliest age possible, helping us, as a society, to train critical consumers who can navigate an increasingly complex informational world.

Acknowledgements

We thank Maggie Fernicola, Renee Johnsson, Tirone Johnson, Katie Hernandez, Karen Levush, Fiona Lachman, and Jason Solinsky for assistance with data collection and coding. This work was supported in part by a grant from the Spencer Foundation to the first author.

References

Bascandziev, I., & Harris, P. L. (2016). The beautiful and the accurate: Are children's selective trust decisions biased? *Journal of Experimental Child Psychology, 152*, 92–105. https://doi.org/10.1016/j.jecp.2016.06.017.

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

Bernard, S., Castelain, T., Mercier, H., Kaufmann, L., Van der Henst, J., et al. (2016). The boss is always right: Preschoolers endorse the testimony of a dominant over that of a subordinate. *Journal of Experimental Child Psychology, 152*, 307–317. https://doi.org/10.1016/j.jecp.2016.08.007.

Birch, S. A., Vauthier, S. A., & Bloom, P. (2008). Three-and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition, 107*, 1018–1034. https://doi.org/10.1016/j.cognitioin.2007.12.008.

Brosseau-Liard, P. E., & Birch, S. A. J. (2011). Epistemic states and traits: Preschoolers appreciate the differential informativeness of situation-specific and person-specific cues to knowledge. *Child Development, 82*, 1788–1796. https://doi.org/10.1111/j.1467-8624.2011.01662.x.

Butler, L. P. (2020). The empirical child? A framework for investigating the development of empirical habits of mind. *Child Development Perspectives, 14*, 34–40. https://doi.org/10.1111/cdep.12354.

Butler, L. P., Schmidt, M. F., Tavassolie, N. S., & Gibbs, H. M. (2018). Children's evaluation of verified and unverified claims. *Journal of Experimental Child Psychology, 176*, 73–83. https://doi.org/10.1016/j.jecp.2018.07.007.

Chen, E. E., Corriveau, K. H., & Harris, P. L. (2013). Children trust a consensus composed of outgroup members—but do not retain that trust. *Child Development, 84*, 269–282. https://doi.org/10.1111/j.1467-8624.2012.01850.x.

Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: Bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior, 33*, 46–56. https://doi.org/10.1016/j.evolhumbehav.2011.05.005.

Clough, P. T. (2010). The case of sociology: Governmentality and methodology. *Critical Inquiry, 36*, 627–641. https://doi.org/10.1086/655203.

Corriveau, K. H., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science, 12*, 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x.

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. https://doi.org/10.1111/cdev.12240.

Corriveau, K. H., Kinzler, K., & Harris, P. L. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology, 49*, 470–479. https://doi.org/10.1037/a0030604.

Corriveau, K. H., Meints, K., & Harris, P. L. (2009). Early tracking of informant accuracy and inaccuracy. *The British Journal of Developmental Psychology, 27*, 331–342. https://doi.org/10.1348/026151008X310229.

Danovitch, J. H., & Keil, F. C. (2004). Should you ask a fisherman or a biologist? Developmental shifts in ways of clustering knowledge. *Child Development, 75*, 918–931. https://doi.org/10.1111/j.1467-8624.2004.00714.x.

Danovitch, J. H., & Keil, F. C. (2007). Choosing between hearts and minds: Children's understanding of moral advisors. *Cognitive Development, 22*, 110–123. https://doi.org/10.1016/j.cogdev.2006.07.001.

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

Fedra, E., & Schmidt, M. F. (2018). Preschoolers understand the moral dimension of factual claims. *Frontiers in Psychology, 9*, 1841. https://doi.org/10.3389/fpsyg.2018.01841.

Fedra, E., & Schmidt, M. F. (2019). Older (but not younger) preschoolers reject incorrect knowledge claims. *The British Journal of Developmental Psychology, 37*, 130–145. https://doi.org/10.1111/bjdp.12264.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science, 337*(6102), 1623–1627. https://doi.org/10.1126/science.1223416.

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

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology, 69*, 251–273. https://doi.org/10.1146/annurev-psych-122216-011710.

Hetherington, C., Hendrickson, C., & Koenig, M. A. (2014). Reducing an in-group bias in preschool children: The impact of moral behavior. *Developmental Science, 17*, 1042–1049. https://doi.org/10.1111/desc.12192.

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. https://doi.org/10.1111/j.1467-9280.2006.01778.x.

Koenig, M. A., Cole, C. A., Meyer, M., Ridge, K. E., Kushnir, T., & Gelman, S. A. (2015). Reasoning about knowledge: Children's evaluations of generality and verifiability. *Cognitive Psychology, 83*, 22–39. https://doi.org/10.1016/j.cogpsych.2015.08.007.

Koerber, S., Sodian, B., Thoermer, C., & Nett, U. (2005). Scientific reasoning in young children: Preschoolers' ability to evaluate covariation evidence. *Swiss Journal of Psychology, 64*, 141–152. https://doi.org/10.1024/1421-0185.64.3.141.

Köksal-Tuncer, Ö., & Sodian, B. (2018). The development of scientific reasoning: Hypothesis testing and argumentation from evidence in young children. *Cognitive Development, 48*, 135–145. https://doi.org/10.1016/j.cogdev.2018.06.011.

Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology, 52*, 423–429. https://doi.org/10.1037/dev0000089.

Köymen, B., Rosenbaum, L., & Tomasello, M. (2014). Reasoning during joint decision-making by preschool peers. *Cognitive Development, 32*, 74–85. https://doi.org/10.1016/j.cogdev.2014.09.001.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science, 9*, 178–181. https://doi.org/10.1111/1467-8721.00088.

Kuhn, D., & Dean, D., Jr (2004). Connecting scientific reasoning and causal inference. *Journal of Cognition and Development, 5*, 261–288. https://doi.org/10.1207/s15327647jcd0502_5.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15*, 309–328. https://doi.org/10.1016/S0885-2014(00)00030-7.

Kuhn, D., Katz, J. B., & Dean, D., Jr (2004). Developing reason. *Thinking & Reasoning, 10*, 197–219. https://doi.org/10.1080/13546780442000015.

Kushnir, T., Vredenburgh, C., & Schneider, L. A. (2013). "Who can help me fix this toy?" the distinction between causal knowledge and word knowledge guides preschoolers' selective requests for information. *Developmental Psychology, 49*, 446. https://doi.org/10.1037/a0031649.

Kushnir, T., Wellman, H. M., & Gelman, S. A. (2008). The role of preschoolers' social understanding in evaluating the informativeness of causal interventions. *Cognition, 107*, 1084–1092. https://doi.org/10.1016/j.cognition.2007.10.004.

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

Michener, H. A., & Burt, M. R. (1975). Components of "authority" as determinants of compliance. *Journal of Personality and Social Psychology, 31*, 606. https://doi.org/10.1037/h0077080.

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology, 49*, 404–418. https://doi.org/10.1037/

a0029500.
Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science, 16*, 385–390. https://doi.org/10.1111/j.0956-7976.2005.01545.x.
Mills, C. M., & Keil, F. C. (2008). Children's developing notions of (im)partiality. *Cognition, 107*, 528–551. https://doi.org/10.1016/j.cognition.2007.11.003.
Mills, C. M., Legare, C. H., Grant, M. G., & Landrum, A. R. (2011). Determining who to question, what to ask, and how much information to ask for: The development of inquiry in young children. *Journal of Experimental Child Psychology, 110*, 539–560. https://doi.org/10.1016/j.jecp.2011.06.003.
Nurmsoo, E., & Robinson, E. J. (2009). Children's trust in previously inaccurate informants who were well or poorly informed: When past errors can be excused. *Child Development, 80*, 23–27. https://doi.org/10.1111/j.1467-8624.2008.01243.x.
Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology, 47*, 116–129. https://doi.org/10.1016/0022-0965(89)90066-0.
Robinson, E. J., Butterfill, S. A., & Nurmsoo, E. (2011). Gaining knowledge via other minds: Children's flexible trust in others as sources of information. *The British Journal of Developmental Psychology, 29*, 961–980. https://doi.org/10.1111/j.2044-835x.2011.02036.x.
Ronfard, S., & Lane, J. D. (2018). Preschoolers continually adjust their epistemic trust based on an informant's ongoing accuracy. *Child Development, 89*, 414–429. https://doi.org/10.1111/cdev.12720.
Ruffman, T., Perner, J., Olson, D. R., & Doherty, M. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis-evidence relation. *Child Development, 64*, 1617–1636. https://doi.org/10.2307/1131459.
Schulz, L. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences, 16*, 382–389. https://doi.org/10.1016/j.tics.2012.06.004.
Shutts, K., Banaji, M. R., & Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Developmental Science, 13*, 599–610. https://doi.org/10.1111/j.1467-7687.2009.00913.x.
Sobel, D. M., & Letourneau, S. M. (2015). Children's developing understanding of what and how they learn. *Journal of Experimental Child Psychology, 132*, 221–229. https://doi.org/10.1016/j.jecp.2015.01.004.
Sobel, D. M., & Letourneau, S. M. (2016). Children's developing knowledge of and reflection about teaching. *Journal of Experimental Child Psychology, 143*, 111–122. https://doi.org/10.1016/j.jecp.2015.10.009.
Sobel, D. M., & Letourneau, S. M. (2018). Preschoolers' understanding of how others learn through action and instruction. *Child Development, 89*, 961–970. https://doi.org/10.1111/cdev.12773.
Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development, 62*, 753–766. https://doi.org/10.2307/1131175.
Stephens, E. C., & Koenig, M. A. (2015). Varieties of testimony: Children's selective learning in semantic versus episodic domains. *Cognition, 137*, 182–188. https://doi.org/10.1016/j.cognition.2015.01.004.
Terrier, N., Bernard, S., Mercier, H., & Clément, F. (2016). Visual access trumps gender in 3-and 4-year-old children's endorsement of testimony. *Journal of Experimental Child Psychology, 146*, 223–230. https://doi.org/10.1016/j.jecp.2016.02.002.
Tomasello, M. (2018). The normative turn in early moral development. *Human Development, 61*, 248–263. https://doi.org/10.1159/000492802.
Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality. *Annual Review of Psychology, 64*, 231–255. https://doi.org/10.1146/annurev-psych-113011-143812.
Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*, 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x.
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. https://doi.org/10.1111/1467-8624.00304.
Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., et al. (2003). *The development of executive function in early childhood. Monographs of the Society for Research in Child Development*i–151.