Journal of Cognition and Development

Routledge Taylor & Francis Group

ISSN: 1524-8372 (Print) 1532-7647 (Online) Journal homepage: https://www.tandfonline.com/loi/hjcd20

Fostering Children's Reasoning about Disagreements through an Inquiry-based Curriculum

Amanda S. Haber, David M. Sobel & Deena Skolnick Weisberg

To cite this article: Amanda S. Haber, David M. Sobel & Deena Skolnick Weisberg (2019): Fostering Children's Reasoning about Disagreements through an Inquiry-based Curriculum, Journal of Cognition and Development

To link to this article: https://doi.org/10.1080/15248372.2019.1639713

Published online: 24 Jul 2019.

JOURNAL OF COGNITION AND DEVELOPMENT
https://doi.org/10.1080/15248372.2019.1639713

Routledge
Taylor & Francis Group

EMPIRICAL ARTICLE

Check for updates

Fostering Children's Reasoning about Disagreements through an Inquiry-based Curriculum

Amanda S. Haber[a], David M. Sobel[b], and Deena Skolnick Weisberg [a,c]

[a]University of Pennsylvania; [b]Brown University; [c]Villanova University

ABSTRACT

We investigated how young children evaluate disagreements between two people and whether formal education affects this capacity. We compared 120 first graders tested during the 2014–2015 academic year, who received a direct instruction-based curriculum, with 112 first graders tested in the same school system during the 2016–2017 academic year, who received an inquiry-based curriculum. All children were given a belief reasoning task that tested their ability to evaluate disagreements about matters of fact, matters of interpretation, and matters of preference. Children's evaluations of disagreements about interpretations or preferences did not differ depending on curriculum. Children who received an inquiry-based curriculum were more likely to resolve disagreements concerning facts correctly than children who received a direct instruction-based curriculum. When asked to justify their responses to disagreements about facts, children who received the inquiry-based curriculum relied more on an examination of the state of the world. We suggest that an inquiry-based curriculum fosters a greater appreciation for how first-hand experiences can create knowledge.

In situations of disagreement, how do children reconcile different beliefs? In some cases, such reconciliation involves consulting objective sources, but in other cases, there is no objective matter of fact about the situation. Children are often faced with the problem of sorting out competing beliefs from different sources, yet even adults struggle with reconciling different points of view (Kuhn, Cheney, & Weinstock, 2000). The present study, part of a larger project about children's developing scientific and metacognitive reasoning, examines how a school curriculum might influence children's ability to resolve disagreements.

Children's developing understanding of beliefs and knowledge

Research into children's understanding of beliefs has examined children's abilities to understand the mental states of others. By age 2, children produce mental state terms such as *think* and *know* (Shatz, Wellman, & Silber, 1983). By age 3, children recognize that different people can hold different beliefs about unknown situations (e.g., Wellman & Bartsch, 1988; Wellman & Liu, 2004). By age 4, children develop a representational conception of belief, which allows them to appreciate that beliefs can

CONTACT Amanda S. Haber haber317@bu.edu Wheelock College of Education & Human Development, Boston University, 621 Commonwealth Avenue, Room 314, Boston, MA 02215

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hjcd.

be false (e.g., Flavell, Mumme, Green, & Flavell, 1992; Gopnik & Astington, 1988; Perner, Leekam, & Wimmer, 1987; Pillow, 1989, 1993).

However, there is more to children's developing understanding of knowledge beside their ability to reason about mismatches between beliefs and reality. Between the ages of 5 and 8, children come to recognize that beliefs have certain recursive qualities (Perner & Wimmer, 1985). Moreover, they learn that one's beliefs causally relate to the thoughts and emotional reactions that one has (Carpendale & Chandler, 1996; Eisbach, 2004; Flavell, Green, & Flavell, 1995; Lagattuta & Wellman, 2001). This kind of epistemological understanding involves more than just understanding false belief; children come to recognize that beliefs can differ based on the interpretation an individual places on the situation and that this process can be independent of objective reality. Children must learn to navigate cases where the conflict is between a single belief and reality, but also cases where the conflict is between different individuals' beliefs given their relations to an objective reality.

Here, we ask children about three types of conflicting beliefs: differences in beliefs based in a matter of fact, a matter of preference, or a matter of interpretation. For the current purposes, facts are true in an objective sense, do not depend on subjective opinions, and can be verified through observation of the world. For example, a possible fact is that a particular ice cream is made with eggs. In comparison with facts, preferences are personal and involve a specific attitude toward an object. For example, the statement *chocolate ice cream is the best flavor* expresses a preference, which does not have the same objective truth conditions. Finally, interpretation-based beliefs represent a more complex combination of objective and subjective factors. For example, consider the statement that *ice cream is in the freezer* in a house that has two freezers. One person could interpret this by believing that ice cream is in the freezer in the kitchen. Another person could interpret this by believing that the ice cream is in the freezer in the basement. Because the initial statement is ambiguous, either interpretation could be correct; without further knowledge, both beliefs are valid.

Previous research indicates that children can differentiate between fact-based and preference-based beliefs. When shown two people who disagree about matters of fact, 5-year-olds judged that only one person could be right; when shown two people who disagree about a matter of preference, these same children judged that both could be right (Heiphetz, Spelke, Harris, & Banaji, 2013, 2014; Walker, Wartenberg, & Winner, 2012). In terms of matters of interpretation, Kuhn et al. (2000) argued that young children struggle with coordinating and judging contradictory but potentially correct statements. In support of this argument, Piekny and Maehler (2013) found that the ability to reason about sets of evidence emerges slowly over the first few years of schooling, and is especially difficult if there is any ambiguity in the evidence. Even adults find various aspects of this kind of belief coordination difficult (Barzilai & Eshet-Alkalai, 2015). Similarly, while 5-year-olds begin to register that an ambiguous figure can have multiple interpretations in their own perception (e.g., Mitroff, Sobel, & Gopnik, 2006), it is not until later in development that children register that different individuals can hold different interpretations of the same ambiguous figure (Beck, Robinson, Ahmed, & Abid, 2011).

Resolving disagreements and curricular instruction

Children's understanding of different and conflicting beliefs begins to mature between the ages of 5 and 7 (Heiphetz et al., 2013), a finding consistent with the literature on the development of an interpretive theory of mind (Carpendale & Chandler, 1996). Because this timing coincides with their entry into formal schooling, different curricula might contribute to how children obtain and evaluate factual knowledge and recognize disagreements when such knowledge is in conflict. We examined first-graders' reasoning about the appropriateness of holding different beliefs, contrasting one cohort of students who received a more traditional curriculum based on direct instruction with a different cohort of students who received a more inquiry-based curriculum. This project was conducted in a single school district which transitioned a direct instruction-based to an inquiry-based curriculum during the project period, retaining the same teachers and the same classrooms.

In a direct instruction-based curriculum, children rely heavily on the teacher to obtain factual knowledge. A large body of previous work suggests that children are quite capable of navigating this process of receiving testimony and judging whom to trust (e.g., Birch, Vauthier, & Bloom, 2008; Corriveau & Harris, 2009a, 2009b; Harris & Corriveau, 2011; Harris, Koenig, Corriveau, & Jaswal, 2018; Jaswal & Neely, 2006; Koenig & Harris, 2005), and traditional classroom instruction takes advantage of this.

In contrast, an inquiry-based learning curriculum emphasizes the idea that children actively construct their own knowledge through exploration, question-asking, self-directed experimentation, and investigation (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011; Edson, 2013). According to the *National Science Educational Standards*, "inquiry is an active learning process – something that students do, not something that is done to them" (Anderson, 2002, p. 2). Through inquiry-based learning, children engage in a process of asking questions, making predictions, investigating, and evaluating evidence, as well as reflecting on their knowledge (Klahr & Nigam, 2004; Seraphin, Philippoff, Kaupp, & Vallin, 2012; White & Frederiksen, 1998).

To illustrate the difference between the two curricula in practice, consider a teacher who wants children to learn about characteristics of living and non-living things. In a direct instruction curriculum, a teacher may begin by explaining the difference between living things (things that grow or have other biological processes) and non-living things (thing that do not have such processes). Next, the teacher may provide the students with three examples of living things and non-living things and explain why each item fits into the given category. The students in this classroom are relying on the teacher to acquire the knowledge for differentiating between living and non-living things.

In contrast, in an inquiry-based learning curriculum, the teacher may begin by telling students that they are going to learn about living and non-living things. Instead of directly telling students the difference between living and non-living things, the teacher may scaffold students learning by engaging them in a structured exploration of non-living and living things in their classroom environment. The teacher might facilitate this exploration by providing students with materials to record living and non-living things that they find in the classroom, or encouraging them to use tools to help them with classifying items (e.g., non-fiction informational texts about living things in the classroom library). In addition, the teacher would inform and encourage students to explain their reasoning for how they classified an item to their classmates and provide evidence to support their decision. Following the activity, the entire class could engage in a discussion about what it means for something to be a living or non-living thing.

While students exposed to both curricula potentially learn about living and non-living things, the process by which this knowledge was acquired looks different. Instead of solely relying on the teacher for information, children who receive an inquiry-based curriculum are obtaining their own information through direct interactions with the world. They are also evaluating evidence, developing arguments and reflecting upon their own knowledge, as they prepare to explain their decisions to their classmates. This process of acquiring knowledge, in contrast with direct instruction, might encourage greater reflection on children's own understanding of how they are learning. Children's experiences with these different methods of learning may thus affect their understanding of the objectivity of knowledge and hence of the ways in which different individuals may disagree.

Classroom instruction tends to incorporate elements of both direct instruction and inquiry-based learning; these are not mutually exclusive constructs. Further, teachers can implement inquiry-based activities in a variety of ways and with more or less scaffolding (Kidman & Casinader, 2017). The current study is thus most accurately described as contrasting a curriculum that placed more emphasis on direct instruction with one that incorporated more inquiry-based learning (see detailed description below).

Overview of the current study

The goal of this investigation is to examine whether these different curricula affect children's ability to resolve disagreement. We administered a measure based on children's epistemic development following Heiphetz et al. (2013; see also Walker et al., 2012). This task shows children two characters, attributes a belief to each character, and asks children to explain if one character or both characters can be right in their beliefs. We built on this design to investigate children's ability to evaluate disagreements between people about matters of fact, matters of interpretation, and matters of preference.

We compared first graders tested during the 2014–2015 academic year with first graders tested during the 2016–2017 academic year in the same school system. The curriculum in that school system shifted from more direct instruction-based learning in 2014–2015 to a focus on inquiry-based learning in 2016–2017. The direct instruction-based curriculum asked children to rely mainly on teachers to acquire knowledge, and students had few opportunities to actively engage in investigations, ask questions, or develop higher-level thinking or cognitive processing skills (analyzing data, arguing beliefs based upon evidence, reflecting upon knowledge). Furthermore, in the direct instruction curriculum, little time was dedicated to learning social studies or science.

The school system adopted a new curriculum based upon the goals of the Common Core State Standards as well as a push to maximize instructional time. This new curriculum, which we categorize as more inquiry-based, integrated content knowledge and process skills and focused heavily on bringing more science content into the classroom, especially in primary grades. This curriculum, aligned with the Pennsylvania State Standards in English Language Arts, Science and Social Studies, was designed to integrate basic literacy skills (reading, writing, speaking) with science and social studies topics. It is structured around asking essential questions, focused on fostering students' critical thinking, reasoning and analytical skills. These essential questions are designed to be open-ended, thought-provoking questions that require high-order thinking (e.g., making predictions, analyzing findings, reflecting upon knowledge) in order to lead students to ask additional questions (McTighe & Wiggins, 2013). For example, for the essential questions "How can we use patterns to explain and predict?" and "How are living things similar and different?", first graders' reading and writing instruction focused on informational and explanatory texts and their science instruction included an inquiry unit on weather and seasons.

To demonstrate this difference concretely, in the direct instruction curriculum, first-graders learned about how plants grow by examining pictures and listening to information from the teacher. In the inquiry-based curriculum, first-grade students engaged in more hands-on activities. They planted seeds, asked questions, observed, developed hypotheses, and reflected upon their initial predictions about plant growth. These students used scientific equipment, such as barometers, as they learned about wind and tracked patterns of the moon and stars. They also were asked to consider how the patterns they were observing emerged over time. These newly designed lessons also integrated science content with non-fiction literature, where teachers would have students read and write in language arts about the content they were learning in science. In addition to spending more time on science, teachers also had more opportunities to discuss issues of facts and opinions with students. In contrast to the direct-instruction curriculum, where students only learned about these issues during literacy time, the inquiry-based learning curriculum used informational texts to connect these concepts to the science units.

Further, the inquiry-based learning curriculum was designed to support teachers with in helping students with the notion of a *productive struggle*. The shift in curriculum forced teachers to move away from "providing students with the answers [towards] helping them to understand that it is really about the process, not necessarily about getting the right or wrong answer" (Curriculum Director, personal communication, January 14, 2019). Through this integrated curriculum, first graders acquired knowledge about the world through their experimentation and investigation of different habitats and natural world phenomenon (e.g., changing seasons). Further, these first graders had the opportunity to gain "a deeper understanding of their world, how culture and nature influence their world, and most importantly, how to use reading and writing to expand their understanding of the world in which they live" (accessed from the school website, September 8, 2018).

Methods

Participants

The final sample included 232 first-graders (M_{age} = 87.09 months, age range: 73.73–103.87 months; 118 female, 114 male) recruited from a suburban school district. The racial distribution of the sample (as identified by parental report) was as follows: 191 were Caucasian, 16 were African American, 13 were Asian, 1 was Native Hawaiian, 4 were other or of mixed descent, and 7 were unknown. One hundred twenty children (M_{age} = 86.92 months, age range: 73.73– 97.36 months; 60 female, 60 male) were tested towards the end of the 2014–2015 academic year (May 4–7, 2015). One hundred twelve children (M_{age} = 87.28 months, age range: 79.93– 103.86 months; 58 female, 54 male) were tested towards the end of the 2016–2017 academic year (May 8–19, 2017).

Materials

Children were shown six laminated cards (8.5 × 11 inches) that depicted pairs of characters, three male pairs and three female pairs (matched to the child's gender). Characters were referred to with gender-neutral names (e.g., Casey, Jessie), so that the same names could be used for all participants.

Three smaller laminated cards (8.5 × 5.5 inches) that each depicted a different shape were used. Children tested in 2015 saw a big black square (3.25 inches per side), a big red square (3.25 inches per side) and a small red square (1.5 inches per side). Children tested in 2017 saw a big yellow circle (3.25 inches in diameter), a big blue triangle (3.25 inches per side), and a small blue triangle (1.75 inches per side). The shapes changed between 2015 and 2017 because in 2017 we collected data both on the group of first-graders reported here and on a longitudinal follow-up of the children tested in 2015 (not reported here). Because some children who participated at the 2017 time point had previously participated in 2015, we wanted to ensure that those children did not respond based on any memory of their previous answers.

We also used a cardboard barrier (approximately 3 feet wide and 18 inches tall) to block the participants' view of the cards and a penny.

Procedure

All children were tested individually in a quiet room at their school, separate from their classroom. All children at both time periods were tested by the same experimenter. Children's testing sessions were recorded (either audio or video) based upon parent permission. Children received a sticker and certificate for their participation.

The testing procedure was based on the procedure used by Heiphetz et al. (2013). At the beginning of the task, the experimenter placed the three shape cards on the table. The experimenter said, "I have some cards with different shapes on them. We're going to hear what my friends think about them." Using this setup, each child was presented with three trial types: *fact, interpretation* and *preference* (order counterbalanced across participants). There were two questions within each trial type. Below, we use the three shapes used by the cohort tested in 2017 (big yellow circle, big blue triangle, and small blue triangle) to describe the procedure.

Fact trial

On this trial, the experimenter introduced the penny and said that she was going to hide it under one of the shape cards. She then put up the barrier so that children could not see the hiding event. She pretended to hide the penny, but actually she put the penny in her lap so that there would be no visual clues as to which card the penny might be under. After this, the barrier was taken down. The experimenter then told the participant that she hid the penny under the big yellow circle. Then, the experimenter placed one picture of two characters in front of the child on the table and attributed a belief to each character: "This is Casey. And this is Jessie. Casey thinks that the penny is under the big yellow circle. Jessie thinks that the penny is under the big blue triangle." Next, the experimenter asked children two questions to determine if each character could be right about which shape the penny is under.

First, she asked "Could Jessie be right about the penny being under the big blue triangle?" After children responded, the experimenter asked children to justify their response. Then, the experimenter asked, "Could Casey be right about the penny being under the big yellow circle?" and again asked the child to justify his or her response. Note that on this trial, the experimenter explicitly said that she hid the penny under the big yellow circle. Therefore, only one of the two characters (Casey) can be correct.

Interpretation trial

This trial was identical to the Fact trial, except that the experimenter stated that she hid the penny under a blue triangle. Because the wording is ambiguous, it was not clear whether the penny was under the big triangle or the small triangle. Then, the experimenter told the child about two new characters' beliefs: "This is Riley. And this is Peyton. Riley thinks the penny is under the small blue triangle. Peyton thinks the penny is under the big blue triangle." Next, the experimenter asked the child if each character could be right about where the penny is and justify his or her responses. Here, both characters could possibly be right due to the ambiguity in the experimenter's information.

Preference trial

The experimenter introduced a pair of characters and told the child about their preferences with respect to the shapes: "This is Adrian. And this is Taylor. Adrian really likes the card with the small blue triangle and Taylor really likes the card with the big yellow circle." Again, children were asked about each character's beliefs: "Could Adrian be right about liking the small blue triangle? Could Taylor be right about liking the big yellow circle?" In both cases, as in the other trial types, the experimenter asked children to justify their responses.

Children also participated in other tasks measuring their diagnostic reasoning abilities and their understanding of science as part of a larger project. Those tasks are not relevant to the current investigation. They will not be discussed here.

Coding

Responses were recorded on-line by the experimenter and by a second researcher who was observing the testing sessions. In addition, a third researcher double-checked the responses after the testing sessions by comparing the researchers' codes to each other and with the video or audio recording. No discrepancies were found.

Children were scored as passing each trial if they responded to the two questions posed on the trial with a particular pattern. For the Fact trial, a correct response involved the child stating that the character who thought the penny was under the big yellow circle was correct and that the character who thought the penny was under the big blue triangle was incorrect. Children were scored as passing this trial if they responded in this way and were scored as not passing if they responded any other way. Note that two children did not answer one of the questions on the Fact trial, and their responses were counted as incorrect. Excluding these children from the analysis does not change any of the reported results below.

For the Interpretation trial, both characters could be right on the questions. Children were scored as passing this trial if they responded in this way and scored as not passing if they responded any other way.

For the Preference trial, children were scored as passing if they said it was OK for both characters to like what they liked or if it was not OK for both characters to not like what they liked. Given that preferences are subjective, it is OK if both characters like the shapes because they can like whatever they want, but it is also OK for both to not like them because liking can't be judged as correct or incorrect (though this pattern only reflected 12% of participant responses). Incorrect responding on the Preference task was only indicated by the child saying one character was right while the other was wrong.

Justification coding

Justifications on the Fact and Interpretation Trials were coded in a similar manner. We categorized responses into three categories: *testimony, world*, and *perceptual* (see Table 1 for example responses). For the *testimony* code, the child referred to information acquired from the experimenter. To receive a testimony code, the child must explicitly have said the words "because you said" or "you told me" in his or her response. For the world code, the child referred to the state of the *world*, some examples include "because it is a blue triangle," "because it was a yellow circle" and "because it is big, round and yellow." For the *perceptual* code, the child referred to something s/he can hear or see about the cards or the penny.

For the Preference Trials, we categorized justifications into three different categories: *opinion, character*, and *subject* (see Table 2 for examples). For the *opinion* code, the child referred to opinions, which are always right, or can't be right or wrong. For the *character* code, the child referred to something about the character to explain why the character likes that card. For the *subject* code, the child expressed his or her own opinion without referring to the character.

We used three additional codes for all three trial types: *interference, irrelevant* and *don't know*. For the *interference* code, the child referred to information that was necessary for a different trial but not for the current one. For the *irrelevant* code, the child referred to something that is not relevant to the task. For the *don't know* code, the child said, "I don't know."

Table 1. Sample fact and interpretation trial responses and justifications.

Category	Fact Trial Justification	Interpretation Trial Justification
Testimony	"Because you said a yellow circle."	"Because you said it is under a triangle."
World	"Because it is yellow and a circle."	"Because it is a blue triangle."
Perceptual	"Because the card is higher up."	"Because the card is lifted up."
Interference	"Because she likes it."	"Because it is her opinion."
Irrelevant	"Because he's smart."	"She knows it is there."
Don't Know	"I don't know."	"I don't know."

Table 2. Sample preference trial justifications.

Category	Preference Trial Justification
Opinion	"Because it's an opinion. There's no right or wrong."
Character	"Because maybe her favorite color is blue."
Subject	"Because yellow is my favorite color."

To measure the reliability of the justification coding scheme, a random sample of 36 responses (2.6% of the total sample) was independently coded by three research assistants, all blind to children's age, gender, and the type of curriculum children received. The three coders had 97.2% agreement on 35 out of the 36 trials (Kappa = .96). Given this agreement, the rest of the sample was coded by one of the three coders.

Results

Responses on individual trials

We analyzed correct responding using a General Linear Mixed Model assuming a Binary Logistic response, treating age, gender, trial type, and curriculum (i.e., year tested) as Fixed Effects, and our order of questions (i.e., the counterbalancing) as a Random Factor. The model we built analyzed all main effects and 2-way interactions. The overall model was significant, $F(14, 676) = 6.52$, $p < .001$. Within the model, the only significant result was the interaction between trial type and curriculum, $F(2, 676) = 8.24$, $p < .001$. This is illustrated in Figure 1, which shows overall performance on the three trial types by year of testing. Performance on the Fact question differed between the two curricula, $\chi^2(1, N = 230) = 25.59$, $p < .001$, Phi = .33. Performance on the Interpretation and Preference questions did not differ between the two curricula.

Looking more closely at the Fact trial, we considered how children responded on the two individual questions in that trial (i.e., whether children judged that the character who expressed the accurate statement was correct and whether children judged that the character who made the inaccurate statement was incorrect). When the character was accurate, 71.7% of the children given the direct instruction curriculum (2015) correctly stated that the character was accurate, while 91.9% of the children given the inquiry-based curriculum (2017) responded this way. This ratio was significantly different across the two curricula, $\chi^2(1, N = 232) = 15.81$, $p < .001$, Phi = .26. When the character was inaccurate, only 49.2% of the children given the direct instruction curriculum (2015) stated that the character was inaccurate, while 82.1% of the children given the inquiry-based curriculum (2017) did so. This was also a significantly different ratio, $\chi^2(1, N = 230) = 27.57$, $p < .001$,

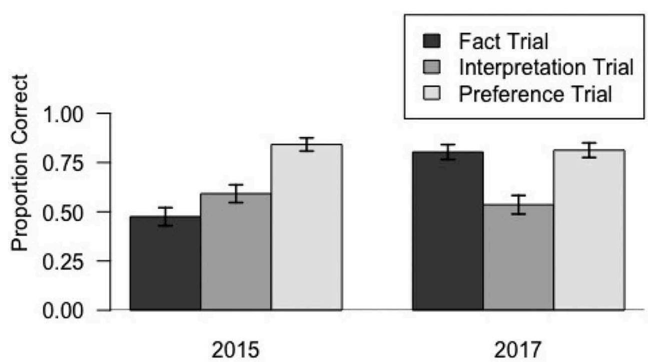

Figure 1. Proportion correct on each trial type by year. The 2015 first graders received a direct instruction-based curriculum and the 2017 first graders received an inquiry-based curriculum. Error bars display standard error.

Phi = .35. Performance by the children tested in 2015 (direct instruction) was no different from chance on this question, Exact Proportions Test, p = .78, while performance on this question by the children tested in 2017 (inquiry) was greater than chance expectations, Exact Proportion Tasks, $p < .05$.

Justification analyses

Children were asked to justify their response to both questions on each trial. Because the coding schemes differed across the trials, we will analyze each separately. For purposes of analysis, we combined the irrelevant, interference, and don't know responses into one category.

Justification of fact trial

For the two questions asked on Fact Trial, we categorized responses into three categories: *testimony, world*, and *perceptual* (see Table 1 for examples). Table 3 shows the percentage of children who provided each type of justification on at least one of the questions on the Fact trials, separated by curriculum and whether children responded correctly on the trial.

We examined whether the type of justification that children generated was related to the curriculum they received, their performance on the Fact trial, and their age. We performed a set of GLMM analyses looking at whether children generated a *testimony, world*, or *perceptual* justification, with trial, age, curriculum and performance as fixed effects. This analysis controlled for within-subject variance because each child received two Fact trials. These results are shown in Table 4. Each model was significant. Age and trial number were unrelated to children's explanations in all of the models. Curriculum and performance on the Fact trial were both related to the types of justifications children generated. Specifically, children who answered the Fact questions correctly were more likely to give *testimony* and *world* justifications than when they were not correct on these questions; the reverse was true for the *perceptual* justifications. However, when controlling for the variance of the other factors, children were more likely to give *testimony* justification when they had received the direct instruction curriculum and more likely to give *world* justifications when they had received the inquiry-based curriculum. Controlling just for correct performance, children given the direct instruction curriculum appealed more to the testimony of others, while children given the inquiry-based curriculum appealed more to evidence in the world. By contrast, children who answered the Fact question incorrectly were more likely to generate *perceptual* justifications or one of the irrelevant justification types.

Table 3. Percentage of trials on which children provided each type of justification on the fact trial, based on curriculum and performance on test trial.

	Testimony	World	Perception
Direct Instruction-Based Curriculum, Incorrect Response (N = 94 trials)	7 (26)	14 (35)	15 (36)
Direct Instruction-Based Curriculum, Correct Response (N = 144 trials)	56 (50)	24 (43)	5 (26)
Inquiry-Based Curriculum, Incorrect Response (N = 29 trials)	7 (26)	3 (19)	21 (41)
Inquiry-Based Curriculum, Correct Response (N = 195 trials)	45 (50)	44 (50)	4 (19)

Notes. Standard deviations shown in parentheses. Numbers for each group do not add up to 100% because we do not present data from children who failed to generate a justification of one of these three types or who said "I don't know". Each child provided two justifications for the Fact Trial.

Table 4. GLMM exploring whether the type of justification that children generated was related to the curriculum they received or to their performance on the fact trial.

	Testimony	World	Perception
Overall Model	*F*(4, 455) = 13.29 *p* < .001	*F*(4, 455) = 8.10, *p* < .001	*F*(4, 455) = 5.06, *p* = .001
Effect of Curriculum Positive beta value indicates odds of inquiry over direct instruction	β = − 0.47 SE = 0.22 *t* = − 2.14 *p* = .033	β = 0.71 SE = 0.23 *t* = 3.13 *p* = .002	β = 0.09 SE = 0.40 *t* = 0.22 *n.s.*
Effect of Performance on Fact Question Positive beta value indicates odds of right over wrong	β = 2.77 SE = 0.39 *t* = 7.10 *p* < .001	β = 1.30 SE = 0.33 *t* = 3.98 *p* < .001	β = − 1.74 SE = 0.41 *t* = − 4.28 *p* < .001
Effect of Age Negative beta value means increasing odds for each month of age	β = − 0.04 SE = 0.02 *t* = − 1.71 *n.s.*	β = 0.01 SE = 0.03 *t* = 0.37 *n.s.*	β = − 0.02 SE = 0.04 *t* = − 0.41 *n.s.*
Effect of Trial Positive beta value indicates odds of second trial over first.	β = − 0.40 SE = 0.22 *t* = − 1.81 *n.s.*	β = − 0.28 SE = 0.22 *t* = − 1.24 *n.s.*	β = 0.34 SE = 0.38 *t* = 0.89 *n.s.*

Table 5. Percentage of trials on which children provided each type of justification on the interpretation trial, based on curriculum and performance on test trial.

	Testimony	World	Perception
Direct Instruction-Based Curriculum, Incorrect Response (N = 55 trials)	5 (23)	20 (40)	20 (40)
Direct Instruction-Based Curriculum, Correct Response (N = 185 trials)	9 (29)	25 (44)	13 (34)
Inquiry-Based Curriculum, Incorrect Response (N = 64 trials)	0 (0)	9 (29)	30 (46)
Inquiry-Based Curriculum, Correct Response (N = 160 trials)	11 (32)	41 (49)	15 (36)

Notes. Standard deviations shown in parentheses. Numbers for each group do not add up to 100% because we do not present data from children who failed to generate a justification of one of these three types or who said "I don't know". Each child provided two justifications for the Interpretation Trial.

Justification of interpretation trial

For the Interpretation Trial, we categorized responses into three categories: *testimony, world*, and *perceptual* (see Table 1 for examples). Table 5 shows the percentage of children who provided each type of justification on at least one of the questions on the Interpretation trials, separated by curriculum and whether children responded correctly on the trial.

We examined whether the type of justification that children generated was related to the curriculum they received, their performance on the Interpretation trial, and their age. As above, we performed a set of GLMM analyses, controlling for within-subject variance because children received two Interpretation trials and looking at whether children generated a *testimony, world*, or *perceptual* justification. Trial, age, curriculum and performance were fixed effects. These results are shown in Table 6. The overall model for the *testimony* justifications did not reach the threshold for statistical significance, and the only significant parameter indicated that children who answered the question correctly were more likely to give a *testimony* justification. The overall models for the *perceptual* and *world* justifications were significant. Children who answered the test question incorrectly were more likely to give a *perceptual* justification, regardless of

Table 6. GLMM exploring whether the type of justification that children generated was related to the curriculum they received or to their performance on the interpretation trial.

	Testimony	World	Perception
Overall Model	$F(4, 457) = 2.28$, $p = .06$	$F(4, 457) = 4.67$, $p = .001$	$F(4, 457) = 3.49$, $p = .008$
Effect of Curriculum Positive beta value indicates odds of inquiry over direct instruction	β = 0.04 SE = 0.36 $t = 0.12$ *n.s.*	β = 0.48 SE = 0.22 $t = 2.21$ $p = .028$	β = 0.29 SE = 0.25 $t = 1.13$ *n.s.*
Effect of Performance on Interpretation Question Positive beta value indicates odds of right over wrong	β = 1.68 SE = 0.67 $t = 2.51$ $p = .013$	β = 1.13 SE = 0.30 $t = 3.83$ $p < .001$	β = − 0.87 SE = 0.27 $t = -3.25$ $p = .001$
Effect of Age Negative beta value means increasing odds for each month of age	β = − 0.07 SE = 0.04 $t = -1.66$ *n.s.*	β = 0.01 SE = 0.03 $t = 0.19$ *n.s.*	β = − 0.03 SE = 0.03 $t = -0.94$ *n.s.*
Effect of Trial Positive beta value indicates odds of second trial over first.	β = − 0.23 SE = 0.38 $t = -0.61$ *n.s.*	β = − 0.24 SE = 0.22 $t = -1.10$ *n.s.*	β = 0.36 SE = 0.26 $t = 1.39$ *n.s.*

any other factor. Generating a *world* justification showed effects of both performance and curriculum, with children receiving the inquiry-based curriculum generating more *world* justifications.

Justification of preference trial

For the Preference Trials, we categorized justifications into three different categories: *character, opinion*, and *subject* (see Table 2 for examples). The frequency of these justifications are shown in Table 7. Given the infrequency of the *subject* code, we did not analyze it further. We examined whether generating a *character* or *opinion* justification was related to the curriculum they received, their performance on the preference trial, and their age via the same GLMM analyses used above. These results are shown in Table 8. Both models were significant, with more children generating *character* justifications when they had received the inquiry-based curriculum and more children generating *opinion* justifications when they had received the direct instruction curriculum.

Table 7. Percentage of trials on which children provided each type of justification on the preference trial, based on curriculum and performance on test trial.

	Character	Opinion	Subject
Direct Instruction-Based Curriculum, Incorrect Response (N = 25 trials)	16 (37)	8 (28)	0 (0)
Direct Instruction-Based Curriculum, Correct Response (N = 215 trials)	45 (50)	32 (47)	1 (6)
Inquiry-Based Curriculum, Incorrect Response (N = 35 trials)	6 (24)	11 (32)	6 (24)
Inquiry-Based Curriculum, Correct Response (N = 189 trials)	66 (48)	18 (39)	1 (7)

Notes. Standard deviations shown in parentheses. Numbers for each group do not add up to 100% because we do not present data from children who failed to generate a justification of one of these three types or who said "I don't know". Each child provided two justifications for the Preference Trial.

Table 8. GLMM exploring whether the type of justification that children generated was related to the curriculum they received or to their performance on the preference trial.

	Character	Opinion
Overall Model	$F(4, 457) = 11.36, p < .001$	$F(4, 457) = 6.79, p < .001$
Effect of Curriculum Positive beta value indicates odds of inquiry over direct instruction	β = 0.80 SE = 0.21 $t = 3.81$ $p < .001$	β = − 0.78 SE = 0.24 $t = -3.30$ $p = .001$
Effect of Performance on Preference Question Positive beta value indicates odds of right over wrong	β = 2.64 SE = 0.47 $t = 5.66$ $p < .001$	β = 1.16 SE = 0.48 $t = 2.43$ $p = .015$
Effect of Age Negative beta value means increasing odds for each month of age	β = 0.02 SE = 0.02 $t = 0.80$ *n.s.*	β = 0.09 SE = 0.03 $t = 3.21$ $p = .001$
Effect of Trial Positive beta value indicates odds of second trial over first.	β = 0.55 SE = 0.21 $t = 2.65$ $p = .008$	β = − 0.31 SE = 0.23 $t = -1.34$ *n.s.*

Discussion

We examined first-graders' ability to evaluate disagreements between two people, taking advantage of a planned change in curriculum implemented by a school district to investigate what impact schooling might have on this ability. The performance of first-graders in 2015, who received a primarily direct-instruction-based curriculum, was contrasted with that of first-graders from the same school and classrooms in 2017, who received a more inquiry-based curriculum. Both groups of children received a measure of their understanding of disagreements over matters of preference, matters of fact, and matters of interpretation.

Matters of preference

There were no differences between the two groups on their understanding of how two individuals could disagree about preferences, and performance on these questions was generally high. This aligns with prior work showing that, by age 5, children understand that different people can hold different preferences without conflict (Heiphetz et al., 2013). However, the two groups of first graders did differ with respect to the type of justification that children used to explain their responses. First-graders in the direct instruction-curriculum generated more *opinion* responses and fewer *character* responses than first-graders in the inquiry-based learning currciulum. One possible reason for this could be that children who received the inquiry-based curriculum may have gained a greater ability to use evidence about the characters' personal beliefs to justify whether they could be correct or incorrect. This idea of supporting one's argument with evidence is more reflective of learning in the inquiry-based learning currcium.

Matters of fact

Judgments about fact-based disagreements did differ between the two groups of first graders, both in terms of overall performance and in the types of justifications children used to explain their responses. First-graders in this sample who had received the inquiry-based curriculum were more likely to understand disagreement about facts than first-graders who had received the direct-instruction curriculum. This understanding requires children to recognize that information generated by others can be false or inconsistent with observed data. Moreover, children who answered correctly were more likely to give a *world* or *testimony* justification when they had received the inquiry-based curriculum. The children who answered correctly were more likely to give only a *testimony* justification when they had received the direct-instruction curriculum, a difference that we discuss in the next section.

Matters of interpretation

Judgments about interpretation-based disagreements did not differ between the two curricula. One possibility for this lack of a difference is that experiences requiring interpretation of beliefs were uncommon in the classroom. Even within the framework of the inquiry-based curriculum, children may have been focused on asking questions about the topics under study but not necessarily on generating answers or on noticing differences among their classmates' answers. We did find, however, that children given the inquiry-based curriculum generated more *world* justifications than children who received the direct-instruction curriculum. This was particularly true when they responded correctly on the measure. This finding is consistent with the hypothesis that inquiry-based curricula afford children with better opportunities to calibrate information generated by others (i.e., their reports on their belief states) with observed data.

Taken together, the results from this study suggest that different school curricula may impact children's abilities to coordinate conflicting beliefs. Although this study explored differences between an inquiry-based learning and direct instruction curriculum, it is important to recognize that there is not a clear dichotomy here, as variation exists even within one type of curriculum. Therefore, future research should pay more attention to the nature of the inquiry-based learning curriculum, which might be at the root of some of these effects found in this study. Although we are unable to address this issue in detail because we did not contrast different types of inquiry-based currciula, the results from the current study nevertheless have implications for two areas of cognitive development research: children's trust in others' testimony and their development of metacognition.

Children's trust in testimony

Children's early factual knowledge largely depends on the testimony of other people (Harris et al., 2018). Previous research indicates that children utilize a variety of cues to determine what information to believe (Harris & Koenig, 2006; Harris et al., 2018; Mills, 2013; Sobel & Kushnir, 2013). For example, young children are more likely to accept information from informants who have been accurate in the past over previously inaccurate informants (e.g., Birch et al., 2008; Pasquini, Corriveau, Koenig, & Harris, 2007) Across a range of situations, by 4 years of age, children tend to favor accurate informants over those with other desirable characteristics including familiarity (Corriveau & Harris, 2009a), age (e.g., Jaswal & Neely, 2006) or accent (e.g., Corriveau, Kinzler, & Harris, 2013). In addition, children attend to the competency of an informant, preferring to learn from an individual who provides a noncircular rather than a circular explanation (e.g., Baum, Danovitch, & Keil, 2008; Corriveau & Kurkul, 2014; Mercier, Bernard, & Clement, 2014).

The present study could be conceptualized as a test of children's understanding of testimony, as children needed to figure out how to interpret an experimenter's statement about two characters' beliefs. Our findings suggest that the type of curriculum children receive in formal schooling may influence when and how children rely on others for information. Specifically, children who often learn through direct instruction may rely more heavily on information from others, because this type of curriculum encourages them to accept the information they are told. Given this, it is reasonable for children to adopt a strategy of treating others' information as factual, even if it might contrast with observed data. This would make it more difficult to assess situations when two people generate contrary information about unambiguous events, as demonstrated by children's generally poorer performance with the Fact trials when they had received a direct-instruction curriculum. This hypothesis is consistent with the view that young children have a "default bias to trust" (Jaswal, Croft, Setia, & Cole, 2010, p. 1541) and that it takes cognitive control to process information in order to assess its veridicality (Jaswal et al., 2014). In contrast, exposure to pedagogical situations in which one discovers information oneself might help children to combat this default bias.

Our data additionally demonstrate that children's pedagogical experiences might affect how they learn to coordinate the information they hear with the data they observe in the world. Because children who learn through inquiry-based experiences play an active role in their learning, they may become more likely to rely on the information they gain through their direct interactions with the world. Such an account is consistent with our finding that children who received the inquiry-based curriculum generated more world-based justifications, particularly for correct answers. These children may have been more likely to recognize that only one informant could be accurate because they could describe how they were assessing the informants' statements.

Metacognition

Metacognitive development is a fundamental component of reasoning in everyday social interactions and scientific thinking, as students "learn how to learn" (Schneider, 2008; White & Frederiksen, 1998). Previous research indicates that students' use of metacognitive strategies had a direct effect on their positive attitudes towards learning science (Jahanangard, Soltani, & Alinejad, 2016; Leopold & Leutner, 2015). As children acquire this metacognitive knowledge, they develop an awareness of their own learning and generate strategies for solving problems and recalling information (Chatzipanteli, Gregoriadis, & Gregoriadis, 2014; White & Frederiksen, 1998).

Beginning in early elementary school, children gain the metacognitive skill of active control over their own cognitive processes (Chatzipanteli et al., 2014; Flavell et al., 1995; Kuhn, 2000; Lagattuta & Wellman, 2001). Further, prior work suggests that opportunities to engage in "collaborative student discourse," where children reflect on what they have learned through class discussions, may foster reasoning and argumentative skills (Mercier, 2011, p. 183). In all of these cases, as in our task, children demonstrate their developing understanding of the possible relations between minds and the world. Through this process, they potentially also acquire a new awareness of their own cognition.

The current work therefore suggests that some prerequisites to this developmental achievement may be fostered by an inquiry-based curriculum, which provides children with the opportunity to "participate in 'doing' science as scientists" (Harris, Phillips, & Penuel, 2012, p. 771). In this study, the school's inquiry-based learning curriculum was centered around essential questions, which integrated content and process to "model the kinds of thinking that students need to emulate and internalize if they are to learn to [think on their own]" (McTighe & Wiggins, 2013, p. 23). Such experiences can allow children to monitor and become more aware of their own learning and reflect upon their knowledge (Alfieri et al., 2011; Martinez, 2012; Tanner, 2012). As a result, this growing metacognitive awareness can facilitate their learning and achievement in school (Kuhn & Pearsall, 1998), including reading and math performance (Schneider, 2008), the capacity to transfer knowledge acquired in one environment to another, and the acquisition of the ability to recognize one's strengths or weaknesses when completing a task (Pintrich, 2002). Tasks like the one used in the current study, which begin to illustrate the need to explicitly navigate multiple points of view, could be helpful in boosting the development of these crucial cognitive skills.

Acknowledgments

This research was supported in part by NSF DRL-1660655 (to DSW) and NSF DRL-1661068 (to DMS). We would like to thank all of the children who participated in this study, as well as their parents, school staff at data collection sites and high school student volunteers who assisted with running the study. Thanks also to the members of the Cognition & Development Lab for their assistance with data collection.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation [DRL-1660655,DRL-1661068].

ORCID

Deena Skolnick Weisberg http://orcid.org/0000-0002-4000-4941

References

Alfieri, L., Brooks, P., Aldrich, N. J., & Tenenbaum, H. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. doi:10.1037/a0021017.supp

Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education*, *13*(1), 1–12. doi:10.1023/A:1015171124982

Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives un comprehension of multiple author viewpoints. *Learning and Instruction*, *36*, 86–103. doi:10.1016/j.learninstruc.2014.12.003

Baum, L. A., Danovitch, J. H., & Keil, F. C. (2008). Children's sensitivity to circular explanations. *Journal of Experimental Child Psychology*, *100*(2), 146–155. doi:10.1016/j.jecp.2007.10.007

Beck, S. R., Robinson, A. N., Ahmed, S., & Abid, R. (2011). Children's understanding that ambiguous figures have multiple interpretations. *European Journal of Developmental Psychology*, *8*(4), 403–422. doi:10.1080/17405629.2010.515885

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*(3), 1018–1034. doi:10.1016/j.cognition.2007.12.008

Carpendale, J. I., & Chandler, M. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development*, *67*(4), 1686–1706. doi:10.2307/1131725

Chatzipanteli, A., Gregoriadis, V. G., & Gregoriadis, A. (2014). Development and evaluation of metacognition in early childhood education. *Early Childhood Development and Care*, *184*(8), 1223–1232. doi:10.1080/03004430.2013.861456

Corriveau, K., & Harris, P. L. (2009a). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*(3), 426–437. doi:10.1111/j.1467-.2008.00792.x

Corriveau, K., & Harris, P. L. (2009b). Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Developmental Science*, *12*(1), 188–193. doi:10.1111/desc.2009.12.issue-1

Corriveau, K. H., Kinzler, K. D., & Harris, P. L. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology*, *49*(3), 470–479. doi:10.1037/a0030604

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*(5), 1827–1835. doi:10.1111/cdev.12240

Edson, M. T. (2013). *Starting with science: Strategies for introducing young children to inquiry*. Portland, ME: Stenhouse.

Eisbach, A. (2004). Children's developing awareness of diversity in people's trains of thought. *Child Development*, *75*(6), 1694–1707. doi:10.1111/j.1467-8624.2004.00810.x

Flavell, J. H., Green, F. L., & Flavell, E. R. (1995). Young children's knowledge about thinking. *Monographs of the Society for Research in Child Development*, *60*(1), 1–113. doi:10.2307/1166124

Flavell, J. H., Mumme, D. L., Green, F. L., & Flavell, E. R. (1992). Young children's understanding of different types of beliefs. *Child Development*, *63*(4), 960–977. doi:10.1111/j.1467-8624.1992.tb01675.x

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*(1), 26–37. doi:10.1111/j.1467-8624.1988.tb03192.x

Harris, C. J., Phillips, R. S., & Penuel, W. R. (2012). Examining teachers' instructional moves aimed at developing students' ideas and questions in learner-centered science classrooms. *Journal of Science Teacher Education*, *23*(7), 769–788. doi:10.1007/s10972-011-9237-0

Harris, P. L., & Corriveau, K. H. (2011). Young children's selective trust in informants. *Philosophical Transactions of the Royal Society Biological Sciences*, *366*(1567), 1179–1187. doi:10.1098/rstb.2010.0321

Harris, P. L., & Koenig, M. (2006). Trust in testimony: How children learn about science and religion. *Child Development*, *77*(3), 505–524. doi:10.1111/j.1467-8624.2006.00886.x

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

Heiphetz, L., Spelke, E. S., Harris, P. L., & Banaji, M. R. (2013). The development of reasoning about beliefs: Fact, preference, and ideology. *Journal of Experimental Social Psychology*, *49*(3), 559–565. doi:10.1016/j.jesp.2012.09.005

Heiphetz, L., Spelke, E. S., Harris, P. L., & Banaji, M. R. (2014). What do different beliefs tell us? An examination of factual, opinion-based, and religious beliefs. *Cognition Development*, *30*(1), 15–29. doi:10.1016/j.cogdev.2013.12.002

Jahanangard, Z., Soltani, A., & Alinejad, M. (2016). Exploring the relationship between metacognition and attitudes towards science of senior secondary students through a structural equation model analysis. *Journal of Baltic Science Education*, *15*(3), 340–349.

Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, *21*(10), 1541–1547. doi:10.1177/0956797610383438

Jaswal, V. K., & Neely, L. A. (2006). Adults do not always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*, *17*(9), 757–758. doi:10.1111/j.1467-9280.2006.01778.x

Jaswal, V. K., Perez-Edgar, K., Kondrad, R. L., Palmquist, C. M., Cole, C. A., & Cole, C. E. (2014). Can't stop believing: Inhibitory control and resistance to misleading testimony. *Developmental Science*, *17*(6), 965–976. doi:10.1111/desc.12187

Kidman, G., & Casinader, N. (2017). *Inquiry-based teaching and learning across disciplines*. London, UK: Palgrave Macmillan UK. doi:10.1057/978-1-137-53463-7

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667. doi:10.1111/j.0956-7976.2004.00737.x

Koenig, M., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*(6), 1261–1277. doi:10.1111/j.1467-8624.2005.00849.x

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*(5), 178–181. doi:10.1111/1467-8721.00088

Kuhn, D., Cheney, M., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15*(3), 309–328. doi:10.1016/S0885-2014(00)00030-7

Kuhn, D., & Pearsall, S. (1998). Relations between metastrategic knowledge and strategic performance. *Cognitive Development*, *13*(2), 227–247. doi:10.1016/S0885-2014(98)90040-5

Lagattuta, K. H., & Wellman, H. M. (2001). Thinking about the past: Early knowledge about links between prior experience, thinking, and emotion. *Child Development*, *72*(1), 82–102. doi:10.1111/1467-8624.00267

Leopold, C., & Leutner, D. (2015). Improving students' science text comprehension through metacognitive self-regulation when applying learning strategies. *Metacognition and Learning*, *10*(3), 313–346. doi:10.1007/s11409-014-9130-2

Martinez, M. E. (2012). What is metacognition? *Phi Delta Kappan*, *87*(9), 696–699. doi:10.1177/003172170608700916

McTighe, J., & Wiggins, G. P. (2013). *Essential questions: Opening doors to student understanding*. Alexandria, VA, USA: ASCD.

Mercier, H. (2011). Reasoning serves argumentation in children. *Cognitive Development*, *26*(3), 177–191. doi:10.1016/j.cogdev.2010.12.001

Mercier, H., Bernard, S., & Clement, 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

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology*, *49*(3), 404–418. doi:10.1037/a0029500

Mitroff, S. R., Sobel, D. M., & Gopnik, A. (2006). Reversing how think about ambigious figure reversals: Spontanenous alternating by uniformed observers. *Perception*, *35*, 709–718. doi:10.1068/p5520

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, *43*(5), 1216–1226. doi:10.1037/0012-1649.43.5.1216

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*(2), 125–137. doi:10.1111/j.2044-835X.1987.tb01048.x

Perner, J., & Wimmer, H. (1985). “John thinks that Mary thinks that … ” Attribution of second-order beliefs by 5- to 10- year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471. doi:10.1016/0022-0965(85)90051-7

Piekny, J., & Maehler, C. (2013). Scientific reasoning in early and middle childhood: The development of domain-general evidence, evaluation, experimentation and hypothesis generation skills. *Bristih Journal of Developmental Psychology*, *31*(2), 153–179. doi:10.1111/j.2044-835X.2012.02082.x

Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology*, *47*(1), 116–129. doi:10.1016/0022-0965(89)90066-0

Pillow, B. H. (1993). Preschool children’s understandung of the relationship between modality of perceptual access and knowledge of perceptual properties. *British Journal of Developmental Psychology*, *11*, 371–389. doi:10.1111/j.2044-835X.1993.tb00610.x

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, *41*(4), 219–225. doi:10.1207/s15430421tip4104_3

Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: Major trends and implications for education. *Mind, Brain and Education*, *2*(3), 114–121. doi:10.1111/j.1751-228X.2008.00041.x

Seraphin, K. D., Philippoff, J., Kaupp, L., & Vallin, L. M. (2012). Metacognition as means to increase the effectiveness of inquiry-based science education. *Science Education International*, *23*(4), 366–382.

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, *14*(3), 301–321. doi:10.1016/0010-0277(83)90008-2

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review*, *120*(4), 779–797. doi:10.1037/a0034191

Tanner, K. D. (2012). Promoting student metacognition. *CBE- Life Sciences Education*, *11*(2), 113–120. doi:10.1187/cbe.12-03-0033

Walker, C. M., Wartenberg, T. E., & Winner, E. (2012). Engagement in philosophical dialogue facilitates children’s reasoning about subjectivity. *Developmental Psychology*, *49*(7), 1338–1347. doi:10.1037/a0029870

Wellman, H. M., & Bartsch, K. (1988). Young children’s reasoning about beliefs. *Cognition*, *30*, 239–277. doi:10.1016/0010-0277(88)90021-2

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523–541. doi:10.1111/j.1467-8624.2004.00691.x

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, *16*(1), 3–118. doi:10.1207/s1532690xci1601_2