Can You Do It?: How Preschoolers Judge Whether Others Have Learned

David M. Sobel

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University

Accepted author version posted online: 09 Nov 2013. Published online: 09 Nov 2013.

To cite this article: Journal of Cognition and Development (2013): Can You Do It?: How Preschoolers Judge Whether Others Have Learned, Journal of Cognition and Development, DOI: 10.1080/15248372.2013.815621

To link to this article: http://dx.doi.org/10.1080/15248372.2013.815621

Disclaimer: This is a version of an unedited manuscript that has been accepted for publication. As a service to authors and researchers we are providing this version of the accepted manuscript (AM). Copyediting, typesetting, and review of the resulting proof will be undertaken on this manuscript before final publication of the Version of Record (VoR). During production and pre-press, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal relate to this version also.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions
ARTICLE

Can you do it?: How preschoolers judge whether others have learned

David M. Sobel

Department of Cognitive, Linguistic, and Psychological Sciences Brown University

Address comments concerning this manuscript to: D. Sobel, CLPS Department, Brown University, Box 1821, Providence, RI, 02912. Email: Dave_Sobel@brown.edu

Abstract

Two experiments investigated how preschoolers judge whether learning has occurred. Experiment 1 showed that 3- and 4-year-olds used an individual’s ability to demonstrate knowledge to judge whether s/he had learned something, regardless of that individual’s claim about whether s/he has learned. Experiment 2 considered whether children responded based on just the character’s demonstrative ability or whether children integrate various pieces of mental state knowledge together to make a judgment about learning. Using a similar procedure, preschoolers were first told that the character claimed to be ignorant, and then that they learned or did not learn a piece of information. In these cases, judgments of learning changed when the characters’ claims and demonstrative abilities conflicted. These results suggest that children’s understanding of learning involves the integration of various pieces of mental state knowledge. This process starts in the preschool years, but these data also suggest that crucial developments are taking place after the age of 4.

Young children are remarkable learners. But, what do young children know about the process of learning? Do young children know how and when learning occurs? Do young
children understand that learning requires that their own or another’s knowledge has changed? Do young children recognize that one’s understanding of learning is potentially fallible – that is, that it can be misguided or incorrect?

The answers to these questions are potentially related to children’s understanding of knowledge and belief, because one cannot know something without having learned it. While there is some research on children’s developing concepts of learning as a mental state (e.g., Bartsch, Horvath, & Estes, 2003; Sobel, Li, & Corriveau, 2007), these investigations do not examine what role children’s understanding of knowledge and belief play in their understanding of learning.

There is, however, much work on children’s understanding of their own and others’ knowledge and beliefs. When we consider the literature on children’s understanding of their own and others’ knowledge, there are good reasons to believe the answers to the questions posed above are all ‘no’. When taught new pieces of knowledge, preschoolers typically claim that they knew it all along (Esbensen, Taylor, & Stoess, 1997; Taylor, Esbensen, & Bennett, 1994) or misunderstand when they have learned a specific piece of information (Tang & Bartsch, 2012; Tang, Bartsch & Nunez, 2007). Although 4-year-olds understand that their knowledge changes, it is not until later in development that they can track how they know that change occurred (e.g., Gopnik & Graf, 1988). Four-year-olds also lack an understanding of the indeterminacy of belief. It is not until approximately age 7 that children understand that different people can interpret the same events differently or that a particular belief state can give way to different belief states in
different people (e.g., Carpendale & Chandler, 1996; Eisbach, 2004; Kuhn, Cheney, & Weinstock, 2000; Lagattuta & Wellman, 2001).

That said, when we consider the literature on children’s appreciation that their beliefs do change, there are potential reasons to think that the answers to the questions posed above are all ‘yes.’ Wellman and Liu (2004) documented children’s developing understanding of knowledge and belief during the preschool years. For instance, 3-year-olds register the difference between their knowing and not knowing a piece of information, that they and another person can believe two different things if neither has access to the truth, and that perceptual access is critical to knowledge (e.g., Hogrefe, Wimmer, & Perner, 1986; Pillow, 1989; Pratt & Bryant, 1990; Wellman & Bartsch, 1989). Children appreciate this knowledge before they understand that others’ beliefs can be false (e.g., Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983). Also developing during this time is children’s knowledge that their own beliefs can change (e.g., Gopnik & Astington, 1988), and the difference between their own knowledge and that of another person’s (Wimmer, Hogrefe, & Perner, 1988).

Based on some of the research described above, Perner (1991) suggested that very young children have a “commonsense theory of knowledge” (p. 146). He argued that most relevant to understanding whether another person knows something is to appreciate whether s/he can demonstrate that knowledge. Knowledge allows a person to complete an action related to that information successfully – if one knows how to sing a song, then one should be able to sing that song successfully (assuming no other factor is preventing
the completion of this action). Perner, however, did not consider how children conceptualize learning as a mental state. A simple extension of his account is that children interpret whether one has learned something as whether one currently possesses that piece of knowledge, as indicated by his/her ability to demonstrate it.

Building on this approach, Sobel, Li, and Corriveau (2007), argued that children develop a concept of learning from their understanding of knowledge and other mental states. Instead of positing learning as just whether an individual has a piece of knowledge, children make judgments about whether learning has occurred based on integrating together various pieces of mental state information. If children only have an individual’s demonstrative abilities, this approach makes identical predictions to the commonsense theory of knowledge approach (i.e., can do = has learned). When children have additional information, however, they weigh that information to form a judgment, as opposed to simply using demonstrative ability. Depending on the nature of that additional information, children’s judgments might differ.

The goal of the present experiments is to map out one way those judgments differ. The two experiments consider cases in which children must judge whether learning has occurred based on a character’s demonstrative ability as well as additional pieces of information. In Experiment 1, we told 3- and 4-year-olds about characters who were being taught a piece of information by a teacher. When the teacher was finished, all of the characters were asked to demonstrate the knowledge with mixed results (i.e., some demonstrated they possessed the knowledge while others did not). Prior to this
demonstration, each character made a claim about whether they learned the material, which turned out to be accurate or inaccurate. In each case, children were asked whether the character learned the information. Experiment 2 replicated the method from Experiment 1, but introduced a third piece of information: that the character stated at the outset that s/he was ignorant of the material, so that the claim about learning could reflect the characters’ belief about a change in their knowledge. In both experiments, we also assessed children’s performance on a standard ‘unexpected contents’ false belief measure (based on Gopnik & Astington, 1988).

The two approaches outlined above make different predictions regarding the overall pattern of data that should result from the two experiments. In Experiment 1, children have to integrate two pieces of information – a claim made by the character about whether s/he learned, which is ultimately proven accurate or inaccurate based on the character’s demonstrative ability. Very young children can recognize that claims can be false (Koenig & Echols, 2003; Lee & Cameron, 2000; Pea, 1982), even if they fail to recognize that such claims might not entail their truth-value (the character could think he has learned something when he has not, i.e., hold a false belief) until later in development (e.g., Wellman, Cross & Watson, 2001). Given this, both the commonsense approach and the integrative one hypothesize that children would discount the claim and respond based on whether the character can demonstrate the information.

That said, the two approaches posit that the way in which children are making this judgment differs. The commonsense theory of knowledge posits that judgments about
learning are based on children’s demonstrative ability, and thus the claim information is discounted. The integrative approach suggests that children infer the likelihood that learning occurred given that the character made the specific claims and had the specific demonstrative abilities. Presumably (and consistent with the commonsense theory), demonstrative ability is more important for this judgment. A character’s claim about his/her knowledge might have some influence, particularly when it disagrees with that character’s demonstrative ability.

That said, claims and demonstrative ability could differ in two ways. When the character claims to have not learned a piece of knowledge that he can demonstrate, the positive demonstration might be most influential for children’s judgments. If the character can do it, there is little reason to consider why s/he claimed the opposite. In contrast, when a character claims to have learned something that he cannot demonstrate, the child might be more likely to consider why the character made that claim. The ability to state that learning did not occur in this case might be dependent on the child’s ability to understand that a claim does not necessarily entail its truth-value (i.e., that it could reflect a false belief). This is why both experiments included a false belief measure. The integrative approach suggests that responses in this situation should correlate with performance on the false belief measure in Experiment 1.

The commonsense and integrative approaches also make different predictions when all of the characters initially state that they do not know the material (Experiment 2). Adding this information should be irrelevant for the commonsense account. The integrative
approach requires that the claim the character makes after the teacher’s actions must take
on additional significance, particularly if it disagrees with the initial claim (i.e., if the
class character initially states she does not know, and then states she has learned). In this case,
children might weigh the second claim more heavily than in Experiment 1, as they should
infer that there must be a reason for this change in the character’s beliefs about his/her
knowledge. If that claim is then shown to be false, children might struggle with
subsequent inferences, more so than in Experiment 1.

EXPERIMENT 1

Three- and 4-year-olds were introduced to a set of child characters in a learning
environment, where a teacher was teaching them a song. Some characters claimed to
have learned the song, other claimed to have not learned. Some characters could sing the
song and knew the words; others did not know the words and could not sing the song.
Each child was read four stories representing the four combinations of claims and
demonstrative ability. Children were asked whether the character learned the song.

Methods

Participants

Twenty 3-year-olds (9 girls, $M = 44.70$ months, $SD = 3.10$ months, range 38-49 months)
and twenty 4-year-olds (10 girls, $M = 55.40$, $SD = 3.22$ months, range 50-62 months)
were recruited from flyers posted in local preschools and a list of hospital births from an
urban area. Two additional 3-year-olds were tested, but were not included in the final
analysis; one was not a native speaker of English, and one refused to participate. Thirty-five of the children were Caucasian, 2 were Hispanic, and 3 were of mixed descent.

Materials

Each story was composed of two pictures drawn on a 14 x 22 cm piece of white paper mounted on blue cardboard. The first picture in each story was of a child character. The second picture was of that character singing (represented by a speech balloon with notes in it) or not singing (represented by an empty speech balloon). Wellman, Hollander and Schult (1996) has suggested that while preschoolers do not spontaneously register an understanding of speech balloons, when accompanied by narrative explaining their content, children at this age have little trouble understanding their meaning. As such, both here and in Experiment 2, speech bubbles were used to describe what the character said, and given the text of the story, we believed that children would understand this material.

In addition, one picture of all four characters and a teacher, drawn on a 28 x 22 cm piece of white paper (also mounted on blue cardboard) was used in the introduction. A Band Aids box that contained crayons or a Crayola crayons box that contained candles was also used.

Procedure

Children were tested with a parent or caregiver present by a male experimenter with whom they were first familiarized. Children were shown an initial picture of four children at school with a teacher. They were told, “Here are a group of children at school. The
teacher is teaching the children how to sing a song. The teacher sings the song and all the children listen. All the children hear the song. Let’s look at what the children are doing.”

Children were then told four stories, one for each child character. Each character made a claim ($C$) about whether he/she learned the song, and then the teacher asked if the character could demonstrate ($D$) that knowledge, which he/she either could or could not do. In the agreement stories, the characters’ claims about learning agreed with their ability to demonstrate the knowledge. One character claimed to have learned the song and could sing it; the other character claimed to have not learned and could not sing it (the $C+/D+$ and $C-/D-$ stories respectively). The other two stories were conflict stories, in which the characters’ claims about learning disagreed with their ability to demonstrate the knowledge. In one, the character claimed to have learned the song, but could not sing it (the $C+/D-$ story); in the other, the character claimed not to have learned the song, but could sing it correctly (the $C-/D+$ story). Appendix A shows the script of the stories read to the children.

After hearing each story, children were asked two control questions: whether the character said s/he learned the song, and whether the character sang the song. The order of these questions was counterbalanced across stories. Children received corrective feedback if they answered incorrectly. After these questions, children were asked the test question: whether the character learned the song. Children received the four stories in one of four quasi-random orders, so that each story appeared in first, second, third, and fourth position in one order.
Children were also given a standard unexpected content false belief task (following Gopnik & Astington, 1988) to examine whether performance on the test questions related to children’s developing false belief capacity. Children were first asked to tell the experimenter the name of someone they played with (usually a friend or sibling, confirmed by the parent/caregiver). Children were then shown a Band Aids box that contained crayons or a Crayola crayons box that contained candles (randomly determined). They were asked the other question: “<Friend/sibling > has never seen this box before. When comes in here, what will think is in the box?” After children responded, the experimenter then asked the self question: “Before I showed you what was in the box, what did you think was in the box?” After children responded, the experimenter asked a control question: “What is really in the box?”

Results

Children required corrective feedback on 18% of the control questions across the four stories. Overall, children required more feedback on these control questions for the conflict questions than the agreement questions (on 29% vs. 7% of the questions, \( t(39) = 5.80, p < .01, \) Cohen’s \( d = 1.23 \)), and 3-year-olds required more feedback than 4-year-olds, \( t(38) = -2.48, p = .01, \) Cohen’s \( d = 0.80 \). The overall number of control questions children answered correctly did not significantly correlate with the number of learning questions they responded to correctly, adjusted \( R^2(38) = 0.017, ns \). The pattern of results does not change considerably if we only consider responses from children when they did not require feedback on either control question (and critically, the effect sizes remain
similar). Thus, we will report the findings from the full dataset. Table 1 shows the percentage of time children responded that the character was learning on the four stories.

On the two stories in which the character demonstrated knowledge, children typically said that the character learned the song. There was no relation with age, \((40) = -0.18, ns\), for both the C+/D+ and C-/D+ stories respectively, and children stated the character learned more often than chance on both of these stories, Binomial Tests, both \(p\)-value < .005. On the stories in which the character failed to demonstrate the knowledge, a different pattern of results emerged. Responses significantly correlated with age for both the C+/D- and C-/D- stories respectively, \((40) = -0.36\) and -0.31, both \(p\)-values < .05. Older children were more likely to say that the character did not learn in both cases. On the C-/D- stories, 4-year-olds stated that the character did not learn at almost ceiling levels, and both 3- and 4-year-olds said that the character did not learn more often than chance, Binomial tests, both \(p\)-values < .05. On the C+/D- stories, 3-year-olds’ responses were at chance levels, Binomial test, \(ns\), while 4-year-olds differed from chance, Binomial test, \(p < .005\).

We contrasted the stories in which the character demonstrated the ability to sing the song as opposed to the ones in which she did not. Responses did not differ between the C+/D+ and C-/D+ stories or between the C+/D- and C-/D- stories for either age group, sign tests, all \(p\)-values > 0.12. In general, when the character could sing the song, children claimed that s/he learned the song, regardless of the character’s claim. Similarly, when the
character could not sing the song, children claimed that s/he did not learn, regardless of
the character’s claim.

On the false belief task, children received a score of 1 for answering each of the other and
self questions correctly. These scores were summed, provided they responded correctly
on the control question (if not, they received an overall false belief score of 0). We
treated responses to the control question here differently from the learning stories. In the
learning stores, the controls were asked to ensure children remembered the story content,
and feedback was provided if children answered incorrectly. In the false belief measure,
the control question serves as a test question and no feedback was given. Responding
incorrectly to the control question here indicated that children did not track the actual
contents of the box, so their other responses would be unreliable.

Three-year-olds had a mean false belief score of 0.65 (out of 2, $SD = 0.75$), and 4-year-
olds had a mean score of 1.25 ($SD = 0.72$), significantly greater, $t(38) = -2.60, p < .05$.
Performance on one of the stories – in which the character claimed to have learned, but
failed to demonstrate the knowledge (C+/D-) – correlated with performance on the false
belief measure, $(40) = -0.38, p < .05$, such that children who performed better on the
false belief measure were more likely to say that the character did not learn. This
correlation was marginally significant when controlling for children’s age, adjusted
$R^2(37) = 0.04, p = .106$.

Discussion
Experiment 1 showed preschoolers judged whether learning occurred based on whether a character could demonstrate the ability, regardless of that character’s claims about their learning. In general, the results are consistent with both a “commonsense theory of knowledge” approach and the more integrative one. As described above, the integrative account does predict the significant correlation with children’s false belief performance observed only in the C+/D- story. One could argue, however, that the commonsense theory would also predict such a correlation, as children who do not register others’ false beliefs would be more confused by this particular story (i.e., why would the character make this claim if she couldn’t do it?).

To differentiate these two approaches more carefully, in Experiment 2, we added a piece of information to each story – that the character claimed not to know the information initially. This information should be irrelevant if children were responding on the commonsense theory (i.e., we should expect similar performance to Experiment 1). The integrative account suggests that children will no longer just use the character’s demonstrative abilities when making judgments about learning, and that children’s understanding of representational change (i.e., their false belief performance) will mediate performance differently than in Experiment 1.

**EXPERIMENT 2**

Experiment 2 replicated the procedure of Experiment 1 with three modifications. First, we changed what was being taught and learned from a song to a set of puzzles. Although we did not think that this manipulation affected performance in Experiment 1, learning a
song could be done incidentally, without the character being aware of their learning. Learning how to solve a set of puzzles (particularly, as opposed to a single puzzle) is potentially more explicit and requires awareness of the character’s part, again potentially putting more emphasis on the claims the characters make about their knowledge. Second, each character started by claiming ignorance of the knowledge. For the integrative account, this added a critical piece of information to each story – that the characters were ignorant, and their positive claim or demonstration revealed a change in their understanding of their capacities. Finally, we added a second test question. In addition to asking whether the character learned the information, we also asked whether the character could demonstrate the knowledge a second time. The goal of including this second test question was to consider how reliable responses to the learning question were – that is, would children judge that if a character had not learned, they could not subsequently demonstrate the knowledge (and vice versa).

**Method**

**Participants**

Twenty 3-year-olds (10 girls, $M = 43.30$ months, $SD = 2.66$ months, range 38-47 months) and twenty 4-year-olds (6 girls, $M = 53.40$, $SD = 2.91$ months, range 48-58 months) were recruited from flyers posted in local preschools, a list of hospital births from an urban area, and from a local children’s museum. Two additional children (one 3-year-olds and one 4-year-old) were tested, but were not included in the final analysis; one because of experimental error, one refused to participate. Three children were Hispanic, all other children were Caucasian.
Materials

Each story was composed of four pictures drawn on a 14 x 22 cm piece of white paper mounted on green cardboard. The first picture in each story was of a child character with a speech bubble that demonstrated they did not know how to solve a puzzle. The second picture was of the child character sitting with a teacher solving a puzzle. The third picture was of the child character sitting with the teacher with a speech bubble that indicated whether the character claimed they did or did not learn how to solve the puzzles. The fourth picture was of the child character sitting alone with a new completed or uncompleted puzzle. Figure 1 shows an example. In addition, one picture of the teacher and all the child characters (shown at the beginning of the procedure) was drawn on a 21 x 28 cm piece of white paper mounted on green cardboard. A crayon box full of candles or a band-aids box full of crayons was used in the false belief task.

Procedure

All children were tested by a male experimenter with their parent or caregiver present. Children were shown the initial picture of the characters at school with a teacher. They were told, “Here’s a teacher and a group of children. They are going to play with the puzzles. Let’s see what happens.” Children were then told four stories. Each story started with the character stating that they did not know how to solve the puzzles. The character and teacher then solved a puzzle together. The character then made a statement about whether he/she learned how to solve the puzzles and whether he/she now knew how to solve the puzzles. The teacher then asked if the character could demonstrate that
knowledge, which the character either could or could not do. As in Experiment 1, children heard four stories in which the character either claimed to learn or not learn how to solve the puzzles, and could or could not demonstrate the knowledge, either in agreement (i.e., C+/D+, C-/D-) or in conflict (i.e., C+/D-, C-/D+). Appendix B shows the scripts of stories read to the children.

After hearing each story, children were asked two control questions: whether the character said he/she learned the puzzles when sitting with the teacher and whether the character solved a puzzle by him/herself. The order of these questions was counterbalanced across stories. Children received corrective feedback if they answered incorrectly. After these questions, children were asked whether the character learned how to solve the puzzles with the teacher. They were then asked to justify their response. Finally, children were asked a second test question, whether the character could solve a new puzzle by him/herself. To control for possible order effects, children received the four stories in one of four quasi-random orders, so that each story appeared first, second, third, or fourth in one order. Finally, children were also given the false belief measure, using the same procedure as Experiment 1.

Results And Discussion

Children required corrective feedback on 18% of the control questions across the four stories. Children required more feedback in the conflict stories than the agreement stories (on 23% vs. 14% of the questions), $t(39) = 2.21, p < .05$, Cohen’s $d = 0.50$. Three-year-olds required more feedback than 4-year-olds, $t(38) = -2.45, p < .05$, Cohen’s $d = 0.79$. 
However, as in Experiment 1, the overall number of control questions children answered correctly did not significantly correlate with the number of learning questions children responded to correctly, adjusted $R^2(38) = -0.01$, $ns$. Accuracy on the two test questions did not differ, all McNemar $\chi^2(1, N = 40)$-values $< 3.36$, all $p$-values $ns$, so for each story, accurate responses on the two test questions were combined to form a score from 0-2 (shown in Table 2).

Preliminary analyses revealed no differences on the test questions between genders, the order in which the control questions were asked, or the order in which the stories were presented. As in Experiment 1, the general pattern of results does not change if we only consider responses from children when they answered both control questions correctly. Thus, we will again report the findings from the full dataset.

Similar to Experiment 1, scores on the agreement stories (C+/D+ and C-/D-) did not significantly correlate with age, both adjusted $R^2(38)$-values $< 0.05$, $ns$, with both 3- and 4-year-olds responding correctly greater than chance levels, both $t(39)$-values $> 2.33$, all $p$-values $< .05$, both Cohen’s $d$-values $> 0.52$. Unlike Experiment 1, there was no significant correlation between age and correct responses to the C+/D- story, adjusted $R^2(38) = -0.025$, $ns$, but there was a significant correlation between age and correct responses to the C-/D+ story, adjusted $R^2(38) = 0.18$, $p < .005$. On the C+/D- story, both 3- and 4-year-olds responded at chance levels, both $t(19)$-values $< 0.30$, both $p$-values $ns$. On the C-/D+ story, 4-year-olds responded corrected more often than chance, $t(19) =$
10.38, \( p < .001 \), while 3-year-olds’ responses were not different from chance, \( t(19) = 1.83, \ ns. \)

As in Experiment 1, we contrasted the two stories in which the character demonstrated knowledge (C+/D+ vs. C-/D+) and the two stories in which the character did not (C+/D-vs. C-/D-). Unlike Experiment 1, children were more likely to score higher in when the claim and demonstrative abilities agreed than conflicted, both when the character could demonstrate the knowledge, \( t(39) = 2.68, p = .01 \), Cohen’s \( d = 0.44 \), and when the character failed to demonstrate the knowledge, \( t(39) = -3.01, p = .005 \), Cohen’s \( d = 0.55 \). That is, it was not the case that whenever the character demonstrated the knowledge, children claimed that they had learned.

Justifications to the learning question were coded into one of six mutually-exclusive categories: (1) I don’t know or no response. (2) That the character said he/she learned. (3) That the character sang the song. (4) That the character knew the words to the song. (5) An alternative mental state (e.g. “She wanted to learn the puzzle.” “Teacher taught him the puzzle.”) (6) An alternative irrelevant justification (e.g., “He learned because he was big”). A research assistant blind to the hypotheses of the experiment coded all of the justifications. A second research assistant, who was also blind to the hypotheses of the experiment, coded four children in each age group (32 out of the total 160 justifications, 20%). Agreement was 94%, and disagreements were resolved through discussion with the author.
Justifications to the learning questions were considered for each individual story. On the C-/D+ story in which the character claimed to have not learned the puzzles, but could solve one by himself, the overall distribution of justifications differed between 3- and 4-year-olds, $\chi^2(5, N = 40) = 11.12, p < .05, \phi = .53$. Four-year-olds generated a higher proportion of justifications in terms of the character’s demonstrations than 3-year-olds, 45% vs. 10% of the time, Fisher’s Exact Test, $p < .05, \phi = .39$. No other story showed any significant results regarding the explanations.

Finally, we examined whether responses to the test questions were related to performance on the false belief task. Three-year-olds had a mean false belief score of 1.05 (out of 2, $SD = 0.81$), and 4-year-olds had a mean score of 1.05 ($SD = 0.81$), not significantly different. Children’s false belief score correlated with responses to only one of the test questions: the C-/D+ story, adjusted $R^2(38) = 0.13, p < .05$. If age is first entered into this regression model, children’s false belief scores still explains a significant amount of the variance, adjusted $R^2(38) = 0.07, p < .05$.

To summarize, when preschoolers were asked for judgments of learning, children used the demonstrative ability when that ability agreed with the characters’ final claim about whether they had learned. When that final claim conflicted with the demonstrative ability, preschoolers struggled to resolve the conflict in the C+/D- case. This is different from the pattern of performance observed in Experiment 1, where they could resolve this conflict by age 4. Moreover, on the C-/D+ stories, younger children – particularly those who struggled understanding others beliefs could not resolve the conflict. Again, this is

---

1 All of the examples come from children’s actual explanations.
different from the pattern of performance observed in Experiment 1, where children judged that this character did learn. Moreover, on this story, when we look at children’s justifications for the learning question, the younger children were less likely to justify their response in terms of the character’s behavior (i.e., their demonstration of the knowledge). This suggests that these younger children might have been influenced more by the claims that the character was making.

**General Discussion**

When a character claimed to learn a piece of information and subsequently demonstrated it, preschoolers judged that the character learned that information. Similarly, when a character claimed to have not learned a piece of information and then could not demonstrate it, preschoolers judged that the character did not learn that information. This was true regardless of whether the character initially claimed that s/he did not know the information (Experiment 2) or not (Experiment 1). Children’s age and performance on an unexpected contents task did not influence responses on these questions; preschoolers were above chance in all cases.

Responses differed when there was a conflict between the claim and demonstrative ability. In Experiment 1, where this was the only information presented, children used the demonstrative ability as the basis for a judgment, although there was some development when the claim was positive, but the demonstration was negative. This was potentially due to children’s developing ability to ignore the claim, as reflected by the significant correlation with the false belief measure.
In Experiment 2, when the character initially claims not to possess the knowledge, the subsequent claim about learning seemed to weigh more heavily. When the claim was positive, but the demonstrative ability was negative, judgments were at chance levels, unlike Experiment 1 where at least 4-year-olds showed above-chance level performance. When the claim was negative, but the demonstrative ability was positive, 3-year-olds were at chance (unlike in Experiment 1, where they responded above chance levels).

One difference between Experiments 1 and 2 is that Experiment 1 asked about learning a song, such that the demonstrative ability reflects the character having learned that particular song, while in Experiment 2, the demonstration (and test questions) involve learning new puzzles. Children might have treated learning each individual puzzle as a different event, so demonstrative ability about one puzzle would not generalize to another. In the song-learning scenario, claims were made specific to the song being taught, so children might be more likely to use the demonstrative ability when making a judgment about learning. It is not clear, however, why this difference predicts the observed pattern of results. An interesting question would be to investigate whether children make different judgments about learning a rote event or a generalizable piece of knowledge (and if so, when in development this distinction emerges). The present data cannot speak to this issue, but an integrative account might suggest that such a subtle difference would only be seen after the preschool years.

How do these results relate to the commonsense theory of knowledge and the integrative
approach to understanding learning? The results of Experiment 1 are generally consistent with both accounts – both predict that judgments of learning should be based on the character’s demonstrative ability, which is what preschoolers appeared to use. The process, however, might be different. The commonsense theory would rely on the positive or negative demonstrative ability. The integrative theory would weigh these two pieces of information against each other, potentially with a default that the demonstrative ability is more important to the judgment. When it is positive, the claim can be discounted more easily. When demonstrative ability is negative, a positive claim (i.e., conflicting information) is resolved based on how easily able the child is at determining the claim is false. This potentially explains the observed correlation between the C+/D- story and the false belief measure in this experiment.

The results of Experiment 2 seem more consistent with the integrative account. When the character initially claims not to know something, and then makes the opposite claim (as in the C+/D- story), that new claim might be weighed more heavily when making a judgment of learning. Children’s chance-level performance here (but not in Experiment 1) reflects that they believe the positive claim has equal merit to the negative demonstration.

This approach can also explain why false belief scores were unrelated to responses to the C-/D+ story in Experiment 1, but were related in Experiment 2. In Experiment 1, because the (positive) demonstrative ability is weighed more heavily, children simply use this information. In Experiment 2, children hear two claims that agree, potentially equating
the amount of conflicting information the child has in this story. The better able to child is to reason that a claim could be false, the better able they might be to conclude that the positive demonstration underlies the judgment of learning.

The integrative account suggested here also relates to various findings beyond learning. In particular, children’s judgments about teaching often appear dependent on children’s developing knowledge of belief and knowledge. For instance, Ziv and Frye (2004) told 3- and 5-year-olds about characters who differed in their ability (i.e., one knew how to read and the other did not) and asked them to predict who would teach the other. They found that both age groups responded that the knowledgeable character would teach the ignorant one. More recently, Ziv, Solomon and Frye (2008) showed that 5-year-olds, but not 3-year-olds judged that a character was not teaching another if the second character learned from watching the first without her knowledge. Similarly, they found that 5-year-olds, but not younger children, judged that a teacher who embedded a teaching activity in a game was trying to teach; the younger children stated the character was trying to play. Finally, they found that 5-year-olds, but not 3-year-olds judged that a teacher would teach students based on the teacher’s beliefs about the student’s knowledge, not the student’s actual knowledge states (see also Strauss, Ziv, & Stein, 2002). All of these findings suggest that the more conflicting mental state knowledge children have to integrate into their judgments, the more difficult the judgment will be. Children do not appear to use a single mental state (i.e., knowledge) to make inferences about teaching. Consistent with arguments made by several researchers that children’s understanding of learning and teaching involve similar processes (e.g., Bartsch et al., 2003; Strauss & Shilony, 1994;
Wellman & Lagattuta, 2004), the present data suggest that children do not base judgments of learning on just a single mental state.

The present results also relate to the developing literature on children’s skepticism (see e.g., Mills & Keil, 2005; Heyman, 2008). For instance, Mills and Keil (2005) suggested that children treat claims differently when they are self-serving and when they are not. In both of the present experiments, children (but particularly 3-year-olds) were better at discounting the claim when it was negative (i.e., not self-serving) than when it was positive. This suggests that children are sensitive to this difference at relatively early ages, potentially earlier than Mills & Keil (2005) believed.

The present studies are not without certain limitations. Both experiments only presented one example of each story type. Follow-up work should consider whether children respond consistently to each type of story. Moreover, the present studies do not consider whether other developmental factors influence the process by which children make these judgments. Certainly, children’s developing language and cognitive control capacities might influence how they interpret the information presented to them. Specifying the role of these broad developmental processes in children’s integrative process is a subject for future investigations.

A final speculation is whether children integrate their knowledge of learning into their broader theory of mind, much in the same way that other mental states, like belief, desire, and pretending are cohesively combined (e.g., Gopnik & Wellman, 1994; Perner, 1991;
Wellman, 1990). Future research might want to investigate whether and how children conceptualize the learning process relates to their developing understanding of mental states more generally. Moreover, some have suggested that elementary-school children’s concepts of learning relate to academic achievement or motivation (see e.g., Li, 2004; Li, Yamamoto, Luo, Batchelor, & Bresnahan, 2010). Understanding how these concepts emerge and how preschoolers’ knowledge might interact with these factors are important ways to extend the current results.

**APPENDIX A**

**Stories Used In Experiment 1**

Note: In all stories, children were asked two control questions, whether the character said he/she learned the story, and whether the character could sing the song. The test question was always the same: whether the character learned the song.

Opening Picture: Here are a bunch of children, and they’re at school, and the teacher is singing a song. The teacher is teaching the song to all of the children. All of the children listen, and all of the children hear the song. So let’s see what the children do, okay? Let’s take a look at what the children do.

Agreement Story 1, Claim/Demo (C+/D+):

Picture 1: This boy, he says, “I learned how to sing the song. I learned how to sing the song.” Now a little while later, the teacher comes and he says, “Do you remember the song that I sang? Can you sing it?”
Picture 2: And look what happens: he tries to sing the song, and he knows all the words. He sings the song, look at that!

Agreement Story 2, No Claim/No Demo (C-/D-):

Picture 1: This boy, he says: “I did not learn the song. I did not learn the song.”

Now a little while later, the teacher comes and he says, “Do you remember the song that I sang? Can you sing it?”

Picture 2: And look what happens: he tries, but he does not know any of the words. He does not sing the song.

Conflict Story 1, Claim/No Demo (C+/D-):

Picture 1: This girl, she says, “I learned how to sing the song. I learned how to sing the song.” Now a little while later, the teacher comes and he says, “Do you remember the song that I sang? Can you sing it?”

Picture 2: And look what happens: this girl tries, but she does not know any of the words. She does not sing the song, look at that!

Conflict Story 2, No Claim/Demo (C-/D+):

Picture 1: This girl says, “I did not learn the song. I did not learn the song.” Now a little while later, the teacher comes and he says, “Do you remember the song that I sang? Can you sing it?”

Picture 2: And look what happens: she tries, and she knows all the words. She sings the song, look at that!
APPENDIX B

Stories Used In Experiment 2

Note: In all stories, children were asked two control questions, whether the character said he/she learned the story with the teacher, and whether the character solved a puzzle by his/herself. The test questions were always the same, whether the character learned the puzzle with the teacher, and whether the character could solve a new puzzle by his/herself.

Opening Picture: Here’s a teacher with a bunch of children. They’re going to play with the puzzles. Let’s see what happens

Agreement Story 1, Claim/Demo:

Picture 1: This girl takes a puzzle and says, “I’ve never done the puzzles before. I don’t know how to do them.”

Picture 2: Teacher sits with her and works with her on the puzzle. They solve it together.

Picture 3: The girl says, “I learned how to do the puzzles. I know how to do them.” Teacher then says, “Can you try another puzzle on your own?”

Picture 4: The girl takes a new puzzle and tries to solve it. Look, she does. She does the puzzle all by herself.

Agreement Story 2, No Claim/No Demo:
Picture 1: This girl takes a puzzle and says, “I’ve never done the puzzles before. I don’t know how to do them.”

Picture 2: Teacher sits with her and works with her on the puzzle. They solve it together.

Picture 3: But the girl says, “I did not learn how to do the puzzles. Teacher did it for me. I don’t know how to do them by myself. Teacher then says, “Can you try another puzzle on your own?”

Picture 4: The girl takes a new puzzle and tries to solve it. Look, she can’t do it. She does not do the puzzle by herself.

Conflict Story #1, Claim/No Demo:

Picture 1: This boy takes a puzzle and says, “I’ve never done the puzzles before. I don’t know how to do them.”

Picture 2: Teacher sits with him and works with him on the puzzle. They solve it together.

Picture 3: The boy says, “I learned how to do the puzzles. I know how to do them.” Teacher then says, “Can you try another puzzle on your own?”

Picture 4: The boy takes a new puzzle and tries to solve it. Look, he can’t do it. He does not do the puzzle by himself.

Conflict Story #2, No Claim/Demo:

Picture 1: This boy takes a puzzle and says, “I’ve never done the puzzles before. I don’t know how to do them.”
Picture 2: Teacher sits with him and works with him on the puzzle. They solve it together.

Picture 3: But the boy says, “I did not learn how to do the puzzles. Teacher did it for me. I don’t know how to do them by myself. Teacher then says, “Can you try another puzzle on your own?”

Picture 4: The boy takes a new puzzle and tries to solve it. Look, he does. He does the puzzle all by himself.

ACKNOWLEDGMENTS

This work was supported by NSF (0518161 and 1223777). I would like to thank all of the parents and children who participated in this research, as well as Esra Aksu, Emily Blumenthal, Claire Cook, Julia Franckh, Loraine Fryer, Sarah Munro, and Philip Parker for help with data collection, analysis, and preparation of this manuscript. Kathleen Corriveau and Jin Li provided me with insightful discussion about this topic.

REFERENCES


Pratt, C., & Bryant P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61*, 973-982.


Table 1 Percentage of Children who said that the Character Learned on each of the Four Stories in Experiment 1.

<table>
<thead>
<tr>
<th>Story Type</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>C+/D+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>90</td>
<td>31</td>
</tr>
<tr>
<td>C-/D+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>80</td>
<td>41</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>70</td>
<td>47</td>
</tr>
<tr>
<td>C+/D-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>45</td>
<td>51</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>C-/D-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>25</td>
<td>44</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>5</td>
<td>22</td>
</tr>
</tbody>
</table>

Notes: C = Claim about learning (+ = yes, - = no), D = Demonstrates knowledge (+ = yes, - = no).
Table 2 Learning Score (out of 2) for Children in Experiment 2 on each of the Four Stories

<table>
<thead>
<tr>
<th>Story Type</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>C+/D+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>1.65</td>
<td>0.59</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>1.95</td>
<td>0.22</td>
</tr>
<tr>
<td>C-/D+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>1.30</td>
<td>0.73</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>1.85</td>
<td>0.37</td>
</tr>
<tr>
<td>C+/D-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>1.00</td>
<td>0.86</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>0.95</td>
<td>0.76</td>
</tr>
<tr>
<td>C-/D-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>1.30</td>
<td>0.66</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>1.45</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Notes: C = Claim about learning (+ = yes, - = no), D = Demonstrates knowledge (+ = yes, - = no).
Figure 1. Examples of the drawings used in the two experiments. This is an example of the agreement story in Experiment 2 in which a girl initially claimed she did not know how to solve the puzzles (picture 1), sat with the teacher and solved a puzzle together (picture 2), claimed she learned how to solve the puzzles and claimed she knew how to solve them now (picture 3) and could solve a new puzzle by herself (picture 4).