To give a fish or to teach how to fish? Children weigh costs and benefits in considering what information to transmit.

Hyowon Gweon (hyora@mit.edu)
Veronica Chu (verochu@mit.edu)
Laura E. Schulz (lschulz@mit.edu)

Department of Brain and Cognitive Sciences, MIT

Abstract

Previous developmental research on pedagogy has focused on children's inferences as learners. Here we look at children's inferences as teachers. We explore the hypothesis that young children consider the goal of the learner and rationally provide evidence that is both informative and cost-efficient. Given a toy with an ambiguous causal structure, children selectively performed costly actions to provide disambiguating evidence only when the learner wanted to know how the toy worked; when the learner only wanted to see the toy's effects, children chose less costly actions. These results suggest that children flexibly modify their behaviors as teachers by considering what learners need to know.

Keywords: cognitive development, pedagogy, social learning, causal learning, pragmatics

Introduction

We constantly communicate with others, learning from them and sharing what we know. A particularly powerful form of communication occurs when one person clearly knows something about the world and wants to share her knowledge, and another person clearly wants to learn what she knows. In a pedagogical context, an informant provides information to a learner to help her learn about the world, and the learner updates her beliefs given information from the informant.

As in any other form of communication, pedagogy is often more than a simple, unidirectional transfer of information. The learner might expect the informant to provide helpful information, and the informant might have a strong motivation to conform to the learner's expectation. Computational models of pedagogical learning have formalized this idea as a set of inferences that mutually constrain one another; a knowledgeable, helpful informant selects data that increase the learner's belief in the correct hypothesis, and an ideal learner rationally updates her belief given data from the informant, with the assumption that the data were selected by a helpful, knowledgeable informant (Shafto, Goodman, & Frank, 2012; Shafto, Goodman, & Griffiths, 2014). This suggests that just as the learner considers the informant's knowledge and intent in pedagogical interactions, the informant also considers what the learner wants in order to select the set of information requisite to her expectations.

Prior developmental work has shown children's receptivity to pedagogically transmitted information.. When learning from a teacher, children draw rational inferences about what is being taught (Bonawitz *et al*., 2011) and decide which informants provide helpful, reliable information (e.g., Koenig, & Harris, 2005; Sabbagh & Baldwin, 2001; Birch, Vautier, & Bloom, 2008; Gweon, Pelton, Konopka, & Schulz, *in press*). Studies suggest that certain kinds of interpretive biases in pedagogical contexts are present even in preverbal infants (e.g., Yoon, Johnson, & Csibra, 2008; Futó, Téglás, Csibra, & Gergely, 2010; see Csibra & Gergely, 2009; 2011 for reviews).

Although there has been relatively less focus on children's ability to teach others, some prior work suggest that even very young children can appropriately communicate what they know. For instance, 12-month-olds pointed more often to an object for an adult who was ignorant of the object's location than for an adult who knew where it was (Liszkowski, Carpenter, & Tomasello, 2008). Furthermore, by age three, children expect that a teacher should teach a skill (e.g., how to sing) to a student who lacks the skill rather than the one who already possesses the skill (Strauss, Ziv, & Stein, 2002), suggesting that children have some understanding of what constitutes good teaching. What has been relatively unexplored in the previous literature is the idea that the cognitive capacities that allow us to be good learners may also make us good at sharing information with others as teachers; just as children rationally infer, as learners, what a teacher is trying to communicate, they might be able to tailor the information they provide as teachers with respect to what the learner knows and what she wants to know.

However, as an informant, knowing what information to provide is not a trivial problem. In many pedagogical interactions, explicit requests for information are either absent, or ambiguous with respect to what the learner wants to know. Imagine a toddler pointing to a light-up toy and asking, "What is that?" Even such explicit requests are rather ill posed, as the informant could generate various behaviors depending on exactly what the learner wants to know: the toddler might want the toy, might want to know the name of the toy, see its cool effects, understand how the effect is generated, etc. Depending on what the child wants, an effective teacher might simply give the toy to the child, label the toy, show its effects by activating the toy, or explain that pressing a hidden button on the toy causes it to light up. Even when the informant is helpful and knowledgeable about the world, the space of possible sets of data she could provide is virtually unlimited.

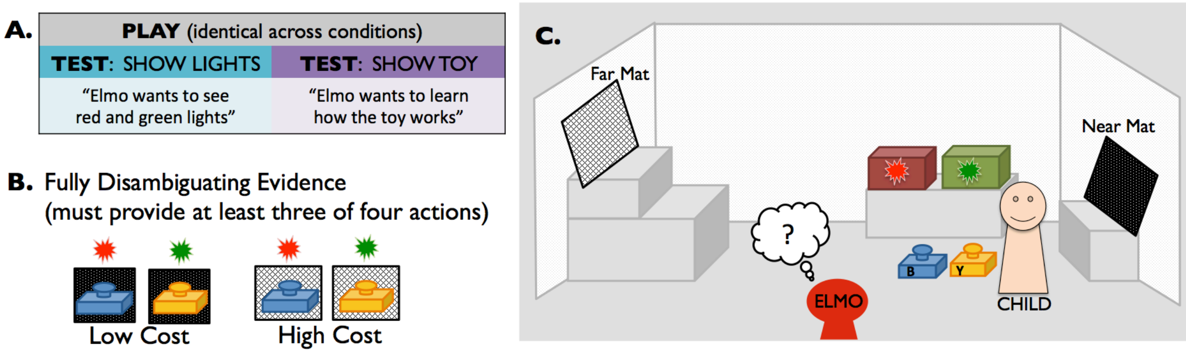

Figure 1. A: Experimental procedure. B: The Blue block on any mat activated the red light, and the Yellow block on any mat activated the green light. To show that blocks, but not mats, affect the color of lights, children had to perform at least 3 of these 4 actions. C: Experimental setup. The Far Mat was much more difficult to reach than the Near Mat, but it was critical to provide disambiguating evidence about the toy.

One way to solve the problem might be to provide as much information as possible. However, information does not come for free, and the amount of transferable data is limited by many factors. For instance, an informant incurs a cost for the time and effort involved in generating the data, and the learner does the same for processing the data. A rational agent should try to minimize the costs of information (e.g., time or effort) while maximizing its benefits (e.g., precision and certainty of our beliefs about the world). Furthermore, not all information is equally useful; its utility depends not only on its truthfulness (the *Maxim of Quality*; Grice, 1975) and necessity for accurate learning (the *Maxim of Quantity*), but also whether the resulting belief is relevant for the current goal of the learner (the *Maxim of Relevance*; see also Wilson & Sperber, 2005). Thus there are two parallel demands, to provide the right *kind* of information, as well as the right *amount* of information. In deciding what information to provide to the learner, it is important to consider what the learner needs and provide just what she needs.

If the learner merely wants to know how to make something happen, you might simply show the target causal relationship. If, however, the learner wants to understand how a toy works, you might give them more elaborated evidence about the causal structure of the toy. Here we explore the hypothesis that even young children, as informants, can (a) infer the right set of evidence the learner needs both in its content and quantity, and (b) incur the cost for generating evidence only when it is necessary for the learner.

To address this question we provided children with a novel causal apparatus (see Figure 1), let them learn the causal relationship themselves, and then asked them to introduce some aspect of the toy to a naïve learner. Imagine an apparatus with two likely potential causes (blocks and mats) and two potential effects (red and green lights); changing the block determines which light will activate, but changing the mat does not. If a learner just wants to see red or green lights, the teacher can simply change the blocks; he has no reason to manipulate the mats (particularly if changing the mat is costly.) By contrast, if the learner wants to know how the toy works, the informant might be most helpful if he showed the learner that changing the block affects the color of the lights, *and* that changing the mat *does not* affect the lights.

Across two conditions, we manipulated whether a naïve agent wanted to see the toy's red and green lights (Show Lights condition) or learn how the toy works (Show Toy condition). We predicted that in comparison to children in the Show Lights condition, children in the Show Toy condition would generate 1) more evidence overall and 2) more informative evidence, even if generating such evidence required children to perform more costly actions.

Experiment

Methods

Subjects Forty-eight children between ages 4 – 6 were recruited from a local children's museum and were randomly assigned to one of two conditions (N=24/condition; mean age in months: 68.5 (Show Lights vs. 69.6 (Show Toy); # boys: 8 (Show Lights) vs. 10 (Show Toy)). Eight children were dropped and replaced due to parental interference (N=2), not completing the procedure (N=2), or experimental error (N=4).

Materials An Elmo hand puppet was used as the naïve learner. The toy consisted of three components: two mats, two blocks, and two boxes; the boxes lit up when the blocks were placed on the mat (see Figure 1 for the experimental setup). Each mat was made of 12" x 12" foam boards. One was colored in black and the other in white, and each mat was also covered with wire mesh with distinct patterns. The black mat was placed right next to where the child sat in the beginning of the experiment (henceforth Near Mat). It was set vertically against a wall in the testing room, approximately 1.5 ft. from the floor. The white mat (Far Mat) was placed vertically near the other side of the wall, on a table approximately 6 ft. from the other mat, and 2.5 ft. from the floor, such that the child had to go around the table or climb on the table to use the mat. Each block was made of acrylic boards, approximately 2 x 2 x 1 inches with a small knob on top. One was colored in blue and the other in yellow. Each light box was approximately 8 x 8 x 4 inches, placed side by side in front of the participant. Each one was covered with red and green felt, respectively. The boxes contained light bulbs visible through a transparent window in front of the box. Each block was wirelessly connected to each box. When the blue block contacted either the white or the black mat, the red light box lit up and stayed lit as long as the block remained in contact with the mat; similarly, the yellow block on either mat activated the green light box. Therefore, even though the mats were necessary for the activation of the lights, the distinction between the two mats was only perceptual and not functionally meaningful for activating different light boxes.

Procedure The experiment took place in a quiet room in the museum. Once the child sat down in front of the red and green light boxes (see Figure 1), the experimenter asked the child to point to each of the light boxes, the yellow and the blue blocks, and the white and the black mats. This ensured that the child saw all components of the toy. Then the experimenter pointed to the blue block and the Near Mat, and said "Hmm, why don't you try putting this blue block on this mat, and see what happens?" Once the participant saw that one of the light boxes lit up, the experimenter said, "I have to go write something down, so why don't you go ahead and play with the toy?" and walked out from the child's line of sight.

After approximately one minute, the experimenter returned to the child and covered up the light boxes so that the child was unable to see which box lit up. She asked the child to turn the red light on, and then to turn the green light on. And then she asked two more test questions. First, she took whichever block the child had just used to turn the green light on, brought it near the opposite mat, and asked "if I put this block on *this* mat, will it turn on the same green light, or the different red light?" Second, she took the other block to bring it over the same mat, and asked, "what if I put *this* block here? Will it turn on the same red (green) light, or the different green (red) light?" These questions were used to assess whether the child had learned the causal structure of the toy during play. If the child could not answer the question or explicitly said, "I don't know", the child was given another minute to play (7 of 48 children played for another minute). These children were asked the same questions after their second play.

Finally, the experimenter brought out her friend "Elmo," a silly monster who knew nothing about these toys. In the Show Lights condition, they were told that Elmo really wanted to see red and green lights, and were asked to "show Elmo red and green lights". In the Show Toy condition, children were told that Elmo really wanted to learn how the toy works, and were asked to "teach Elmo how the toy works". Children demonstrated the toy to Elmo for as long as they wanted; when the child said "I'm done", or when they indicated that they were done by putting down the blocks or stopped to look at the experimenter, the experimenter asked "Are you done?" to confirm that the children were indeed done showing Elmo about the toy.

Video Coding Video recordings of the testing sessions were coded using a video annotation software (VCode; Hagedorn, Hailpern, & Karahalios, 2008) by a trained coder blind to condition manipulation. Its outputs were then analyzed using a custom script in MATLAB separately for the initial play with the toy (Play) and during the child's demonstration of the toy to Elmo (Show). For both Play and Show, we coded for each time the child placed a block (Blue, Yellow) on a mat (Near, Far) to turn on a light box; each of these instances was coded as an "Action" (e.g., if both blocks were placed on the same mat, they were coded as two Actions). We also coded each time the child moved from one end of the room to the other end to use a different mat (coded as a "Transition").

We also measured the informativeness of children's during the Show phase in the following two ways. First, we looked at whether or not the child produced disambiguating evidence about the causal structure of the toy during the entire Show duration. For instance, if the child placed each and block by itself on each mat at least once (four actions total), or produced three of these four actions, the child had produced disambiguating evidence about the toy, allowing the learner to see that the blocks determine the light colors, not mats. Second, we analyzed the informativeness of the first four actions during the Show phase and scored them from 0 to 3 (0: fully confounded evidence, 1: disambiguate either the mats or the blocks, by trying one block on each mat or trying each block on one mat; 2: three of four disambiguating actions; 3: all four disambiguating actions).

Results

Children were given identical instructions and questions until the Show phase began. Thus, during the Play phase, we did not predict any differences in how children played with the toy. However, during the Show phase, we predicted that children would produce different behaviors depending on what they were instructed to do. In the Show Lights condition, children were asked to show Elmo red and green lights. The easiest way to do this is to use the Blue and the Yellow blocks on the Near Mat; there is no need to move from one side to the other side of the room to use both the Near and the Far mat. By contrast, in the Show Toy condition, children were asked to teach Elmo how the toy works. In this case, it is helpful to show that the two blocks are causally responsible for the activation of different lights, as well as that the two mats are identical and not causally relevant for determining which light would turn on. Therefore, during the Show phase, we predicted that children in the Toy condition would produce more Actions and Transitions than children in the Lights condition, and be more likely to provide causally informative evidence about the toy.

During the Play phase, children in the Show Lights condition and Show Toy condition did not show differences

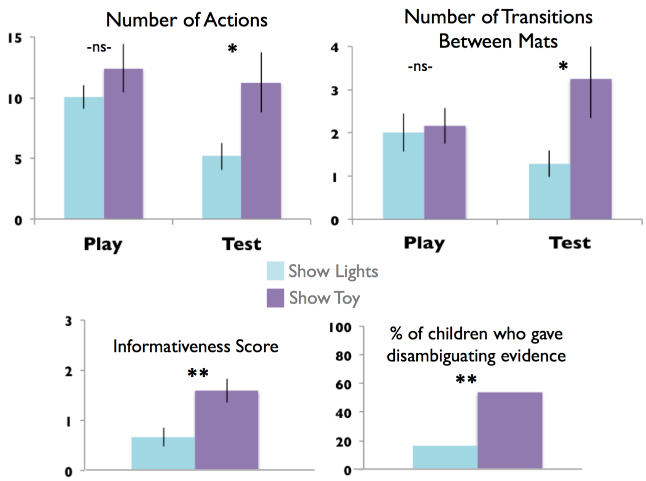

Figure 2. Experiment Results.

in the total playtime (Show Lights vs. Show Toy: 76 vs. 68 sec., $t(46) = 1.05$ $p = 0.30$), number of Actions (Show Lights vs. Show Toy: 10.0 vs. 12.4, $t(46) = 1.08$, $p = 0.29$), or the number of Transitions between the two mats (Show Lights vs. Show Toy: 2.0 vs. 2.2, $t(46) = 0.27$, $p = 0.79$). Children in the two conditions were also equally good at answering the test questions about toy in both conditions. All children used different blocks to activate different lights. 62.5% of children in both conditions understood that the same block on a different mat would activate the same light, and 79.1% (Show Lights) and 70.8% (Show Toy) answered that changing the block would activate a different light.

During the Show phase, even though the duration of demonstrations did not differ significantly across conditions (Show Lights vs. Show Toy: 46 vs. 57 sec., $t(46) = 1.55$, $p = 0.13$), children's behaviors differed by condition (see Figure 2). As predicted, children in the Toy condition produced more Actions (Show Lights vs. Show Toy: 5.1 vs. 11.3, $t(46) = 2.26$, $p = 0.029$) and more Transitions (Show Lights vs. Show Toy: 1.29 vs. 3.25, $t(46) = 2.04$, $p = 0.047$). Further analysis revealed that the number of actions on the Near Mat did not differ across conditions (Show Lights vs. Show Toy: 3.75 vs. 6.37, $t(46) = 1.34$, $p = 0.187$). Instead, the overall difference in the action frequency was driven by the number of actions on the Far Mat (Show Lights vs. Show Toy: 1.42 vs. 4.88, $t(46) = 2.42$, $p = 0.022$) which was necessary only if the children wanted to show Elmo that the mats do not determine the color of lights.

One possibility is that these differences are due to children in the Show Toy condition who failed to understand the causal structure of the toy during the Play phase. Even though children in both conditions demonstrated equivalent knowledge about the toy, given instruction to teach Elmo about how the toy works, children in the Show Toy condition might have produced more diverse actions simply in the hope that this would help Elmo learn. To address this possibility, we split the children in the Show Toy condition into two groups: those who answered all questions correctly (Pass: N=13) and those did not (Fail: N=11). There was no difference in the average number of actions (Pass vs. Fail: 11.5 vs. 11.0, $t(22) = 0.09$, $p = 0.9$) or in the number of transitions (Pass vs. Fail: 3.5 vs. 2.9, $t(22) = 0.33$, $p = 0.74$). Thus children who might have not fully understood how the toy works provided just as many actions and transitions as children who fully understood the toy.

This suggests that children in the Toy condition not only produced more actions but also more costly actions. These costly actions were informative: only acting on the far mat could disambiguate the causal structure of the toy. Indeed, more children in the Toy condition than the Lights condition produced causally disambiguating evidence about the toy (Show Lights vs. Show Toy: 16.6% vs. 54.2%, $p = 0.014$, Fisher's Exact). Interestingly, children were also likely to provide the disambiguating information immediately. The first four actions of children in the Toy condition were significantly more informative than those of children in the Lights condition (Show Lights vs. Show Toy: 0.67 vs. 1.58, $t(46) = 3.01$, $p = 0.004$).

Discussion

Children in our study selectively performed costly actions to generate causally disambiguating evidence only when it is required to fulfill the learner's goals. When children were just showing Elmo the lights, they did so by performing low-cost actions; when Elmo wanted to learn how the toy worked, children not only performed more actions but also actions that were both more costly and more informative.

These results suggest that children, as informants, understand what information to provide to a learner based on his goals. Children in the Show Lights condition generated evidence that was easy to generate (i.e., because it was on the near mat) and failed to fully disambiguate the causal structure of the toy (because they never showed that changing the mat failed to affect the outcome), but still helped Elmo by showing him what he wanted to see. Children in the Show Toy condition generated evidence that was harder to generate (because they moved more often to the far mat) and did disambiguate the causal structure of the toy, and thus provided Elmo with what he wanted to know. Thus children, as teachers rationally balanced the informativeness of evidence and the cost for generating such evidence to generate data that had the highest utility for the learner.

Arguably, given that children's possible actions were naturally constrained by the causal structure of the toy (i.e., there were only four possible pairings of a block and a mat), children might have produced more informative evidence simply by virtue of doing more things. Perhaps children in the Show Toy condition recognized that it was a complicated question and simply did everything they could think of. However, we think this is unlikely. Children could have easily performed other kinds of actions rather than specifically causally disambiguating actions (e.g., they could have slid both blocks on the same mat back and forth, they could have stacked the blocks, etc.). However, children in the Show Toy condition not only produced more causally relevant actions overall, but also produced them immediately after the instruction to teach. Second, the difference between the two conditions emerged from children's actions on the Far Mat rather than on the Closer Mat. This suggests that children in the Show Toy condition did not simply do more actions overall, but that their actions were targeted to produce more causally informative evidence.

By better understanding what others want, we can make better decisions about what to do for the benefit of others. In real life, there are cases where a simple transfer of factual information might suffice, while there are cases worth going through elaborate efforts to derive an abstract understanding of the world in the learner's mind. An old Chinese proverb captures this idea: *Give a man a fish and you feed him for a day, teach a man to fish and you feed him for a lifetime*. Although we often have a sense that the latter kind of teaching is the most worthwhile, sometimes it is simply more efficient to give someone a fish. The ability to flexibly trade costs (e.g., time and effort) for benefits of pedagogy is an important aspect of effective teaching.

In this study, we provided children with an instruction set about what the learner wanted to learn. Children were told either that Elmo wanted merely to see the lights or wanted to learn the causal structure of the toy. In real life however, good informants might not only consider learners' explicit requests about what they would like to learn, but also predict what they would like to learn, or even draw normative decisions about what the learner *ought to learn*. One interesting possibility is that the decision about what to teach and what information to provide will involve a calculation of the learner's expected utility from the data given the learner's mental states such as his beliefs and desires. For example, when a belief inferred from a set of data is likely to be useful repeatedly, an informant might be more willing to teach such data than when the belief might be transiently useful. Just as the ability to learn from information provided by others, the ability to teach others might involve an intuitive understanding of others' knowledge, beliefs, and desires (Theory of Mind), as well as an ability to consider the expected reward and costs of information (i.e., a naïve utility calculus; Jara-Ettinger, Gweon, Tenenbaum, & Schulz, *in press*). Future work might further explore these ideas and shed light on the cognitive mechanisms that underlies our ability to learn from others and share information with others.

The current results provide the first steps in understanding our ability to decide how to efficiently generate information best suited to a learner. Given explicit information about the learner's goal, young children rationally select the right set of evidence for the learner by carefully weighing its cost and informativeness.

Acknowledgments

We thank Julia Ellermeier for help with data collection and Sidney Strauss for helpful comments on an earlier version of this paper. This work was supported by the Center for Brains, Minds, and Machines (CBMM), funded by NSF STC awards CCF-1231216.

References

Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition*, *118*(1), 84–93.

Birch, S. A. J., 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.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–330.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*(3), 341–349.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149–1157.

Futo, J., Téglás, E., Csibra, G., & Gergely, G. (2010). Communicative Function Demonstration induces kind-based artifact representation in preverbal infants. *Cognition*, *117*(1), 1–8.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics: Vol 3, Speech Acts* (pp. 41-58). New York: Academic Press.

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107*(20), 9066-9071.

Gweon, H., Pelton, H., & Schulz, L.E. (2011). Adults and school-aged children accurately evaluate sins of omission in pedagogical contexts. *Proceedings of the 33rd Annual conference of the Cognitive Science Society*, 1242-1247.

Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (*in press*). Sins of omission: Children selectively explore when teachers are under-informative. *Cognition*.

Hagedorn, J., Hailpern, H., & and Karahalios, K.G. (2008). VCode and VData: Illustrating a new framework for supporting the video annotation workflow. *Extended Abstracts of AVI 2008.*

Horn, L. R., & Bayer, S. (1984). Short-circuited implicature: A negative contribution. *Linguistics and Philosophy, 7*(4), 397–414.

Jara-Ettinger, Gweon, Tenenbaum, & Schulz (*in press*). I'd do anything for a cookie (but I won't do that): Children's understanding of the costs and rewards underlying rational action. *Proceedings of the 36th Annual Conference of the Cognitive Science Society.*

Koenig, M., & Harris, P. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*(6), 1261-1277.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, *108*(3), 732–739.

Sabbagh, M., & Baldwin, D. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development, 72*(4), 1054-1070.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning From Others: The Consequences of Psychological Reasoning for Human Learning. *Perspectives on Psychological Science, 7*(4), 341–351.

Shafto, P., Goodman, N.D., & Griffiths, T.L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, *71*, 55-89.

Strauss, S., Ziv, M., & Stein, A. (2002). Teaching as a natural cognition and its relations to preschoolers' developing theory of mind. *Cognitive Development*, *17*(3), 1473–1487.

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences*, *105*(36), 13690–13695.