Young children infer and manage what others think about them

Mika Asaba[a,b,1] and Hyowon Gweon[a,1]

Edited by Renée Baillargeon, University of Illinois at Urbana–Champaign, Champaign, IL; received August 4, 2021; accepted June 3, 2022

We care about what others think of us and often try to present ourselves in a good light. What cognitive capacities underlie our ability to think (or even worry) about reputation, and how do these concerns manifest as strategic self-presentational behaviors? Even though the tendency to modify one's behaviors in the presence of others emerges early in life, the degree to which these behaviors reflect a rich understanding of what others think about the self has remained an open question. Bridging prior work on reputation management, communication, and theory of mind development in early childhood, here we investigate young children's ability to infer and revise others' mental representation of the self. Across four experiments, we find that 3- and 4-y-old children's decisions about to whom to communicate (Experiment 1), what to communicate (Experiments 2 and 3), and which joint activity to engage in with a partner (Experiment 4) are systematically influenced by the partner's observations of the children's own past performance. Children in these studies chose to present self-relevant information selectively and strategically when it could revise the partner's outdated, negative representation of the self. Extending research on children's ability to engage in informative communication, these results demonstrate the sophistication of early self-presentational behaviors: Even young children can draw rich inferences about what others think of them and communicate self-relevant information to revise these representations.

theory of mind | reputation management | communication | social cognition

Significance

Questions about the self are central in our lives: Who am I? Can I do this? Just as common, yet arguably more preoccupying, are questions about the self in the eyes of others: Who do you think I am? Do you think I can do this? While these thoughts might come naturally to adults, what underlies them is a remarkable cognitive feat: the ability to reason about what others think of the self. Even early in life, self-presentational behaviors in humans reflect more than the desire to impress others. Rather, children can make nuanced inferences about what others think about them and strategically communicate to revise others' negative, outdated impressions of them.

Imagine a young child saying, "Watch this!" and proudly tying her shoes in front of her parents. Why would she do this? Given that even young children can help others learn by communicating information about the world (1), one might wonder if she is genuinely trying to teach her parents how to tie their shoes. Yet, here's another possibility: Rather than communicating something about the world, could it be that the child is communicating something about the self, showing off her ability to tie her shoes all on her own?

As humans, we are deeply concerned about our reputations. What (we think) others think about us—particularly, about our internal qualities such as competence or fairness—can have powerful influences on how we learn and perform (2, 3), how we interact with others (4, 5), and even how we think about ourselves (6, 7). As such, we also try to impress others and actively manage our reputation (8, 9). This tendency emerges remarkably early in life (5, 10); beyond being sensitive to others' attention to the self (11), even 3-y-olds cheat more after being told they have a reputation for being "smart" (12) and wait longer in a delay-of-gratification task when they are told that someone (e.g., their teacher or peer) would find out how long they waited (13). By age 5, children engage in prosocial behaviors and avoid antisocial actions when someone else can see their actions (14, 15), and cheat less when they are told that they have a reputation for being "good" (16).

Despite the ubiquity and early emergence of reputational concerns, however, much remains unknown about their cognitive underpinnings: What are the representations and inferential processes that allow such concerns to manifest as strategic self-presentational behaviors? As adults, many of us appreciate how our behaviors can shape what others think of us; beyond being concerned about our reputation, we can also strategically plan our behaviors depending on what others currently think about us and how we want them to think of us. This flexible, nuanced reputation management requires the ability to reason about how others' observations give rise to their current representations of the self, and how additional observations might change those representations. Thus, there are at least two cognitive prerequisites that are critical to turning reputational concerns into strategic self-presentational behaviors: 1) the ability to represent and reason about unobservable mental states—a capacity often referred to as theory of mind (e.g., refs. 17 and 18)—and 2) the ability to use such reasoning to change others' mental states by engaging in flexible, informative communication (1).

Decades of research have investigated the development of theory of mind and communication in early childhood. First, to measure children's ability to reason about others' mental states, researchers have typically manipulated an agent's prior observations and

Author affiliations: [a]Department of Psychology, Stanford University, Stanford, CA 94305; and [b]Department of Psychology, Yale University, New Haven, CT 06511

Author contributions: M.A. and H.G. designed research; M.A. performed research and analyzed data under the supervision of H.G.; and M.A. and H.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: mika.asaba@yale.edu or gweon@stanford.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2105642119/-/DCSupplemental.

Published August 5, 2022.

asked whether children can explicitly predict or explain the agent's behaviors given those observations. For instance, a classic false belief scenario asks: If Sally placed her ball inside a box but then, unbeknownst to her, the ball was moved to a basket, where would Sally look for her ball? Such tasks probe children's ability to infer Sally's representation of the ball's location given her past observation, and understand that this representation is now inconsistent with the ball's actual location. In addition to these standard false belief scenarios, a range of tasks assess children's understanding of desire, knowledge, and emotional states (e.g., ref. 19). The precise timeline of theory of mind development still remains a topic of debate (20–23), especially with regard to implicit measures such as looking time. Nonetheless, children generally succeed at various versions of more "explicit" false belief tasks by age 4 (18, 24), suggesting that they have an abstract, causal understanding of the relationship between others' observations and their resulting mental states.

Second, prior work also suggests that young children have sophisticated communicative abilities that capitalize on rich mental-state inferences (1). Throughout early childhood (particularly during the preschool years), children become increasingly capable of flexibly choosing what information to communicate based on what a learner knows and wants (25–28) as well as what information would be rewarding or costly for the learner to acquire [i.e., expected utility of learning (29, 30)]. Thus, rather than indiscriminately providing any information that is new or interesting, preschool-aged children consider others' knowledge states and selectively communicate information that would help others acquire or update their representations of the world.

Taken together, prior work suggests that, by the end of the preschool years, children can use others' prior observations to infer their mental states (18) and communicate informative evidence to change those mental states (1). Critically, however, existing empirical work on theory of mind development has focused primarily on how children reason about others' knowledge or beliefs about concrete, observable states of the physical world [e.g., the locations or identities of objects (17–24, 31, 32); although, see ref. 33 for recent work on representations of false beliefs about others' emotions]. Likewise, studies that reveal children's abilities as helpful communicators have also focused on how children communicate about the external world [e.g., causal mechanisms or properties of objects (25–30)]. Thus, there are open questions concerning whether these inferential and communicative abilities extend beyond reasoning about others' representations of physical states. In particular, we know relatively little about whether children can represent what others think of them and how such abilities support children's self-presentational behaviors and reputation management more broadly.

Although theory of mind has often been considered to be a candidate cognitive capacity that underlies reputation management (5, 10), little research has directly investigated whether children can represent and manage what others think of them. Rather than manipulating others' prior observations and measuring children's subsequent responses [as in classic theory of mind tasks (18)], prior work in reputation management has typically manipulated the presence of a third-party observer or the observer's properties (e.g., identity or preference), and measured whether children behave more positively when others are watching (e.g., share more stickers), and more so in front of some people than others (11, 13–15, 34). Although one could interpret these findings as reflecting children's representations of what others think of them, it remains possible that children are just motivated to please others and behave more positively around those they want to impress. Other studies have directly manipulated whether children were explicitly told that they have a particular reputation (e.g., for being "nice" or "smart") before engaging in a task (12, 16); in these tasks, children could have simply tried to conform to those expectations by engaging in or avoiding behaviors (e.g., cheating) within the context of the task (particularly when others are watching), rather than representing and attempting to revise what others think of them. Collectively, although these studies have successfully elicited self-presentational behaviors even in young children, they do not provide clear evidence that children, at this age, can represent what others think of them and deliberately communicate self-relevant information to revise or improve these representations.

Building on prior developmental research on theory of mind, communication, and reputation management, we designed our tasks to elicit self-presentational behaviors in ways that reflect young children's mental state inferences and communication about the self. Among many representations that others could have about the self, we focus on those that concern one's competence, a quality that is marked by experiences that are salient even to young children: their performance outcomes [i.e., failures and successes, or quality of their work (35–37)]. For instance, if a child repeatedly fails to activate a toy in front of an adult, and succeeds only after the adult leaves the room, would the child understand that the adult still thinks the child cannot make the toy go? Insofar as the child infers that the adult has an outdated representation of them, the child may attempt to revise it, for instance, by successfully activating the toy when the adult returns ("Look, I can make it go!").

Across four experiments, we targeted 3- and 4-y-old children (with the exception of Experiment 2a, which also included 5-y-olds). The earliest evidence for self-presentational behaviors comes from this age group (13, 16); children at this age also show marked improvements in their performance in standard false-belief tasks (18, 24) as well as tasks that tap into their abilities as helpful communicators (1). In Experiment 1, we provide an initial test of children's ability to revise others' prior observations of the self, by asking to whom they want to communicate about the self. In Experiments 2 and 3, we systematically vary an adult's observations of the world (i.e., how a toy works) and of the child (i.e., whether the child fails or succeeds to activate a toy) to ask how children navigate between two competing goals: communicating about the world (teaching) versus communicating about the self (managing reputation). Finally, in Experiment 4, we ask whether children are more likely to revise others' outdated representation of the self when it is negative (i.e., deflated) than when it is positive (i.e., inflated). Scripts for procedures, data, and analyses for all experiments are available here: https://osf.io/uvmzy/.

Experiment 1

In Experiment 1 ($n = 70$, 3- and 4-y-olds), we test whether young children can track others' observations of their failures and successes to rationally decide to whom to communicate their success. Prior to the main procedure, all children were introduced to a causal device with two buttons (a "toy"; see *Materials and Methods*) by a main experimenter who then left the room. The remaining procedure was conducted by two additional experimenters (Failure Observer and Success Observer) who had no prior interactions with participants and played with them in separate phases.

First, the Failure Observer entered the room and explained how the toy works. She successfully activated the toy to play music, but then the toy did not work when the child tried to

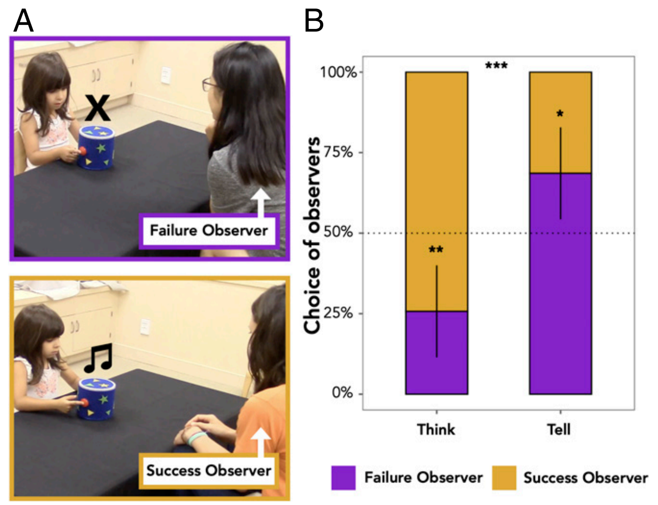

Fig. 1. Methods and results for Experiment 1. (*A*) (*Top*) A participant fails to make the toy work in front of the Failure Observer. (*Bottom*) A participant succeeds in front of the Success Observer. (*B*) Proportion of children's choice of observers when asked "Who thinks you can make the toy go?" (Think) or "Who should I tell that you can make the toy go?" (Tell). Error bars represent 95% CIs; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

activate it.* This sequence was repeated again, and then the Failure Observer left the room. Second, a different experimenter (Success Observer) entered the room; the Success Observer's instructions and interactions with the device were identical to the Failure Observer's, except that the child now succeeded both times. Thus, participants failed twice then succeeded twice on the toy, but the Failure Observer only saw their failures, and the Success Observer only saw their successes (Fig. 1). Finally, the main experimenter returned to the room with a photo of the Failure Observer and a photo of the Success Observer, and asked, "Who thinks you can make the toy go?" (Think Condition) or "Who should I tell that you can make the toy go?" (Tell Condition).†

All children also received an additional memory check question ("Who came in first to play?") to assess whether they could explicitly recall who was watching them during their failures and successes. Correctly answering the memory check question was one of our predetermined exclusion criteria (*Materials and Methods* and *SI Appendix*).‡

Results. As predicted, children were more likely to choose the Failure Observer in the Tell Condition (68.6%) than in the Think Condition (25.7%), $p < 0.001$, Fisher's exact. When asked which Observer thought that they could make the toy go (Think), children selectively chose the Success Observer (74.3%, $p = 0.006$, binomial). When asked to whom to tell their success (Tell), children selectively chose the Failure Observer (68.6%, $p = 0.041$, binomial) (Fig. 1). We did not find evidence for age-related differences in children's choices (*SI Appendix*). These results provide initial evidence that 3- and 4-y-old children readily track others' past observations of their own failures and successes to attribute representations of the self accordingly (Think) and make systematic choices about to whom to communicate their success (Tell).

Although children accurately identified who thought they could make the toy go, their motivation for choosing to tell the Failure Observer about their success remains rather ambiguous: While it is possible that children wanted to change what the Failure Observer thought about them, it is also possible that children were simply motivated to avoid providing redundant information (success) to the Success Observer. We address this issue in Experiment 2, where children interacted with a single observer and were given a chance to successfully demonstrate one of two toys.

Experiment 2

In Experiment 2, we ask whether children can reason about others' representation of the self to strategically navigate between two communicative goals: providing information about the world (i.e., showing a novel toy) or about the self (i.e., showing that they can activate a toy). Given that this task could be more challenging than Experiment 1, we recruited a wider age range in an initial experiment (Experiment 2a: $n = 70$, 3- to 5-y-olds), followed by a direct replication with a larger sample of our primary target age group (Experiment 2b: $n = 100$, 3- and 4-y-olds).

First, the main experimenter brought out two causal devices (toys): a red toy with two buttons and a green toy with two levers (see *Materials and Methods*). A second experimenter (Observer "Anne") then entered the room and stated her ignorance about the toys by saying, "I've never seen these toys before, I don't know how they work!" Then the main experimenter began the Observed Toy Phase with one toy (e.g., the red toy in Fig. 2; toy counterbalanced) and put away the other toy.

In the Observed Toy Phase, the experimenter successfully activated the red toy and gave the child a chance to try, but the child failed to activate the toy. The Observer was in the room watching and acknowledged the experimenter's success ("Oh wow, the toy lights up!") and the child's failure ("Hmm…"). After repeating this sequence once more, the experimenter explained to the child that the two buttons (or levers) have to be pressed at the exact same time for the toy to work and gave the child a third chance to try. Finally, the child succeeded on this third attempt. The critical difference between conditions was when the Observer left the room: In the Absent Condition, the Observer left the room immediately after the experimenter's explanation but before the child's success, whereas, in the Present Condition, the Observer left the room after seeing the child's success.

Next, the Unobserved Toy Phase involved the same procedure with the other toy (e.g., the green toy), but in the absence of the Observer; the experimenter was always successful, while the child initially failed twice and succeeded on the third attempt (Fig. 2).

Finally, at test, the experimenter asked, "Now you can show Anne [the Observer] one of these toys. Which toy do you want to show her?" This test question left the goal of communication deliberately ambiguous, allowing children to choose to provide information about the toys (i.e., teach) or information about their abilities (i.e., show off). At the end, children received a memory check question (predetermined exclusion criteria; *Materials and Methods* and *SI Appendix*): "Was Anne watching when you were playing with this toy or this toy?"

Children in both conditions had good reasons to choose the Unobserved (green) Toy. By doing so, they could provide novel information to the Observer; the Observer had never seen this toy's effect, did not know how it worked, and had not seen the child succeed on it, whereas she had seen the Observed (red)

*In reality, the toy was rigged such that the child's failures and successes were surreptitiously controlled by a hidden remote. This was true for all the toys used in Experiments 1 through 3.

†Prior work suggests that 3- and 4-y-olds understand the meaning of the word "think" (38).

‡All of our experiments had similar memory check questions, and the number of children excluded is reported in *Materials and Methods*. Given the relatively large number of children who did not answer correctly, we also report analyses without this exclusion in *SI Appendix*; the results remain qualitatively the same for Experiments 1 and 2, but are weaker for Experiments 3 and 4 (although in a similar direction).

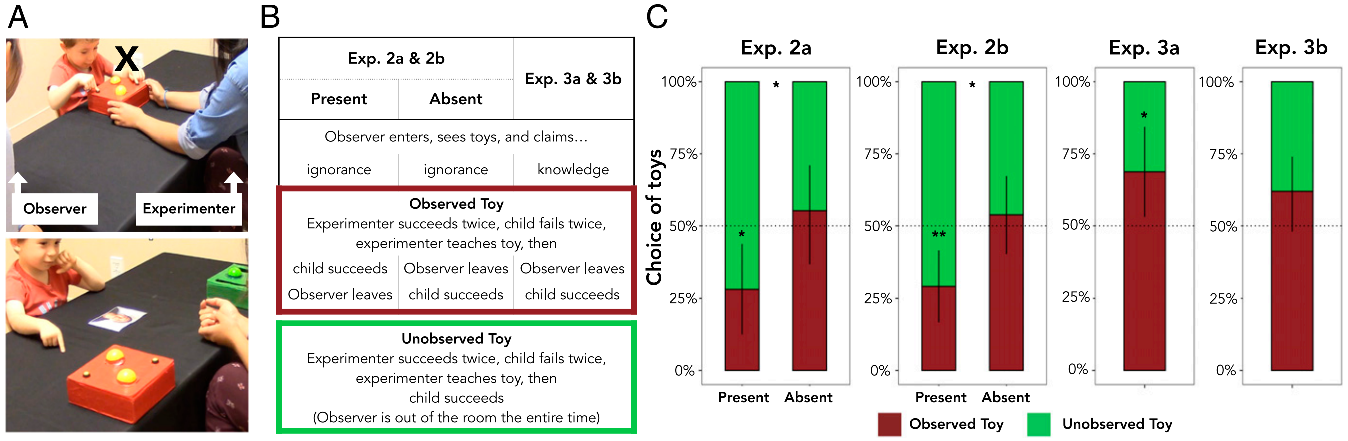

Fig. 2. Methods and results for Experiments 2 and 3. (*A*) (*Top*) A participant attempts to make the Observed Toy go in front of the Observer and the experimenter. (*Bottom*) A participant chooses a toy to show the Observer. (*B*) Schematic of study design. (*C*) Proportion of children's choices for the Observed or Unobserved Toy in each condition. Error bars represent 95% CIs; ** $p < 0.01$, * $p <= 0.05$.

Toy's effect and also heard the experimenter's explanation about how it works. Critically, however, only the children in the Absent Condition had reasons to demonstrate the Observed Toy instead, despite the Observer's familiarity with the toy; by showing their success, children could revise the Observer's past observations of their failure. Thus, our main prediction was a difference between the two conditions; children would be more likely to choose the Observed Toy in the Absent Condition compared to the Present Condition.

Results. As predicted, children in Experiment 2a (3- to 5-y-olds) chose the Observed Toy more often in the Absent Condition than in the Present Condition (55.3% vs. 28.1%; $p = 0.029$, Fisher's exact). Children in the Present Condition selectively chose the Unobserved Toy ($p = 0.020$, binomial), whereas children in the Absent Condition did not show a clear preference ($p = 0.627$, binomial). We found a modest effect of age, suggesting older children were more likely to show a difference between conditions (*SI Appendix*).

We observed a strikingly similar pattern in Experiment 2b, which only included 3- and 4-y-olds. Children were more likely to choose the Observed Toy in the Absent Condition than in the Present Condition (53.8% vs. 29.2%, $p = 0.015$, Fisher's exact); children in the Present Condition strongly favored the Unobserved Toy ($p = 0.006$, binomial), whereas children in the Absent Condition were split between the two toys ($p = 0.678$, binomial). We did not find age-related differences in participant's choices (*SI Appendix*).

These results suggest that, although children in Experiment 2 were generally motivated to provide new information (Unobserved Toy), those in the Absent Condition were more likely to demonstrate their ability to activate the Observed Toy (which was already familiar to the Observer) compared to those in the Present Condition. This tendency is particularly remarkable given that the Observer had already seen the toy and its effect multiple times, and even heard the Experimenter teach the child how it works. Unlike Experiment 1, these responses cannot be explained by a tendency to avoid redundancy. They also cannot be explained by a simple heuristic to engage in positive behaviors or generate positive effects (activating either toy would be positive), or a learned association between success and praise (showing either toy would presumably lead to praise). While children were split between the two toys in the Absent Condition, it is unlikely that they were confused and randomly chose a toy; children in the Present Condition showed a clear preference for the Unobserved Toy (showing that they understood the procedure and the test question), and the two conditions were minimally different in terms of their procedures and demands. Instead, one plausible interpretation of these results is that, even though children in both conditions were driven by the goal to provide new information about the world (i.e., teach how the Unobserved Toy works), only those in the Absent Condition had a competing communicative goal to provide information about the self, specifically because they inferred that the Observer had an outdated representation of their ability to activate the Observed Toy.

An even more stringent test for this interpretation would be to remove the goal to communicate about the world while retaining the goal to communicate about the self: If the Observer already knows how the toys work, which obviates the need to teach the Unobserved Toy, would children in the Absent condition now preferentially choose the Observed Toy to show off their competence? Experiment 3 explores this possibility.

Experiment 3

Experiment 3 consisted of a single-condition, preregistered experiment (Experiment 3a: $n = 32$, 3- and 4-y-olds) and a preregistered direct replication with a larger sample (Experiment 3b: $n = 50$, 3- and 4-y-olds).

The procedure was almost identical to the Absent Condition as in Experiment 2; the only difference was that, instead of stating her ignorance, the Observer (Anne) explicitly stated her knowledge about both toys ("I really like playing with both of these toys. I know how to make both of them go!") and demonstrated that she could successfully make both of them work. Thus, children no longer had a reason to teach the Observer about either toy. In addition, we changed the test question to clarify that the purpose of communication was to tell the Observer about the child's success on either one of the toys. The experimenter said, "I can tell Anne [the Observer] that you can make one of these toys go. Which toy do you want me to tell her that you can make go?" followed by the same memory check question as in Experiment 2 (predetermined exclusion criteria; *Materials and Methods* and *SI Appendix*). Note that choosing either toy could still communicate positive information about self, because the Observer never saw the child successfully activate either toy. However, if children understand that the Observer has an outdated representation of their ability to activate the Observed Toy, they should preferentially choose this toy to repair this representation.

Results. As predicted, children in Experiment 3a preferentially chose the Observed Toy (68.8%, $p = 0.050$, binomial, preregistered analysis). A similar trend was found in Experiment 3b but with a weaker, nonsignificant effect (62%, $p = 0.119$, binomial, preregistered analysis); children were nonetheless more likely to choose the Observed Toy than children in the Present Condition in Experiment 2b (62% vs. 29%, $P = 0.001$, Fisher's exact, preregistered analysis). An exploratory analysis collapsing across Experiments 3a and 3b ($n = 82$ participants total) suggested that children were more likely to choose the Observed Toy (64.6%, $P = 0.011$, binomial).[§] The effect of age as a continuous predictor was not significant in Experiment 3a or 3b, or even in the combined sample (*SI Appendix*).

These results provide further suggestive evidence that children in the Absent Condition wanted to demonstrate their competence on the Observed Toy. Compared to Experiment 2 where children had two competing goals, removing the goal to teach new information about the toys may have allowed the other goal (i.e., the goal to demonstrate their competence) to manifest more clearly. Given the relatively small effect sizes, however, we are cautious to draw strong conclusions. In fact, although Experiment 3 procedures were superficially similar to Experiment 2, there are reasons to believe that this task was more challenging for young children. First, minor changes to the procedure may have increased the overall processing demands of the task (e.g., the Observer's successes on both toys, declaration of her knowledge about both toys, longer test question). Second, children likely needed more sophisticated mental-state reasoning and communicative abilities in order to show a preference for the Observed Toy. Specifically, children had to understand that the Observer's representation of them contained information about which toy they had failed on (the Observed Toy), and that showing the Observed Toy is a more effective way to revise this representation than showing the Unobserved Toy. However, if children were unable to track which toy the Observer saw them fail on, and/or if they simply wanted to show a success on any toy, then choosing either toy would be sufficient. We return to this point in *General Discussion*.

Collectively, our results from Experiments 1 through 3 demonstrate that preschool-aged children track others' observations of their performance and communicate information to revise others' outdated representations of the self. In this sense, children's communication about the self resembles their communication about the world: When others' representations are outdated or false, they provide information that can update or revise those representations.

Yet, a distinctive aspect of self-relevant communication is that we often want these representations to be not just accurate, but also positive. Thus, as prior work on early reputation management suggests (12–16), children's tendency to communicate information about the self may also be influenced by whether that information could improve or damage their reputation. Experiment 4 was designed to test this particular feature of self-relevant communication: Rather than attempt to revise any inaccurate representations of the self, we hypothesized that children may be more likely to revise representations that are deflated (inaccurate and negative) than those that are inflated (inaccurate but positive).

Experiment 4

In Experiment 4 (preregistered, $n = 64$, 3- and 4-y-olds), we used a new, simplified task design. First, rather than having children do two different tasks (i.e., learn how two different toys work), now, children just did one activity—drawing. Prior work suggests that even young children can evaluate different qualities of drawings and are motivated to draw well (e.g., refs. 37, 39, and 40). Second, rather than asking children to directly communicate with the Observer, we asked children which activity they wanted to pursue with the Observer. Critically, the adult Observer was always mistaken (i.e., had an inaccurate representation) about the child's drawing ability, but in a way that was positive (Better Condition) or negative (Worse Condition; random assignment). We asked whether children would be more likely to choose an activity that shows their true competence when the Observer has an inaccurate and negative representation of the self than when the Observer has an inaccurate but positive representation.

First, the participant interacted with the main experimenter in a warm-up task; the child was shown two drawings of trees that differed in quality ("good" and "bad" drawing; Fig. 3) and was asked which drawing is better (predetermined exclusion criteria). The child then made their own drawing of a tree. Next, just as the main experimenter placed it on a pile of other drawings (on the floor, out of the child's view), an Observer came into the room. The experimenter told her, "Hey, [child's name] just made a drawing!" and the Observer said, "Oh, let me take a look!" and picked up a drawing from the pile. The Observer looked at the drawing for 2 s to 3 s, leading the child to believe that the Observer was looking at the child's own drawing, although children could not see the actual drawing. In reality, the Observer picked up a blank sheet of paper (rather than the child's actual drawing), making the Observer blind to condition. The Observer then said "Thanks for showing me!" and placed it back on the pile, and left the room.

Critically, the main experimenter then picked up the "same" drawing that the Observer was looking at, and said, "Oh, this is the drawing that Anne was looking at." In the Better Condition, the experimenter revealed the good drawing from the warm-up task. In the Worse Condition, the experimenter revealed the bad drawing. In reality, the experimenter had surreptitiously placed either the good or the bad drawing on top of the pile of drawings before revealing it to participants. Thus, in both conditions, children were led to believe that the Observer thought she was looking at the child's own drawing, but later learned that the

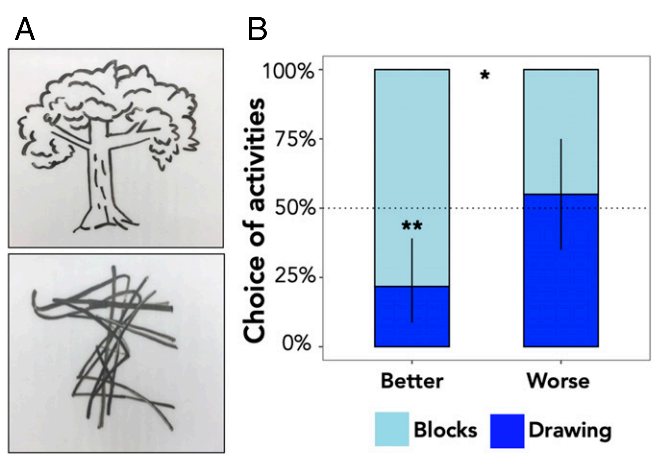

Fig. 3. Stimuli and results for Experiment 4. (*A*) (*Top*) Good drawing used in Better Condition; (*Bottom*) bad drawing used in Worse Condition; good and bad drawings were also used in warm-up task in both conditions. (*B*) Results after excluding participants who made drawings that were rated to be worse than the bad drawing (Worse Condition) and better than the good drawing (Better Condition). Error bars represent 95% CIs; * $p < 0.05$. ** $p = 0.01$.

[§]Note that the results from the combined data should be interpreted with caution, as Experiments 3a and 3b were conducted separately.

Observer had actually looked at someone else's drawing that was either better or worse than their own. Then the Observer reentered the room and brought out a piece of paper and marker, as well as a container with blocks. She asked children the critical question, "Now we can do some drawing or play with blocks. Which one do you want to do?"[¶] After participants engaged in their chosen activity for about 1 min with the Observer, the Observer left the room.

Finally, the main experimenter brought out the good drawing, the bad drawing, and the participant's drawing, and asked a memory check question (preregistered exclusion criteria): "Which drawing was Anne looking at earlier?" The experimenter then replaced the child's drawing with a previous participant's drawing and asked, "Which is the best drawing?" and "Which is the worst drawing?" of the three drawings. By asking children to evaluate another child's drawing in a yoked manner, we were able to independently verify whether children produced drawings that were actually evaluated to be better than the "bad" drawing (Worse Condition) or worse than the "good" drawing (Better Condition) by their peers.

The Observer's representation of the child was false in both conditions; she not only looked at the wrong drawing (i.e., not the child's), but it also differed in quality with respect to the child's own. What differed between conditions was the valence of the Observer's representation. Looking at a drawing that is better than the child's own (Better Condition) would presumably form a positive, desirable representation of the child's performance (or their drawing abilities more generally), whereas looking at a drawing that is worse than the child's own (Worse Condition) would presumably form a negative, undesirable representation. If children can represent these valenced contents of the Observer's representation of the self, they might choose different activities, depending on the condition. Thus, our prediction was a difference across conditions: Children would be more likely to continue drawing (rather than play with blocks) in the Worse Condition than in the Better Condition. Although the drawing task could revise the Observer's representation to reflect their true competence in both conditions, only in the Worse condition would it improve this representation.

Results. Despite a trend in the predicted direction, the difference between the two conditions was not significant (percent choice for drawing: 41.7% [Worse] vs. 25% [Better], $p = 0.193$, Fisher's exact, preregistered main analysis). Preregistered follow-up analyses revealed that children in the Better Condition preferred blocks over drawing ($p = 0.013$, binomial), while children in the Worse Condition did not show a clear preference between blocks and drawing ($p = 0.405$, binomial). Logistic regression with condition and age (continuous) revealed a weak interaction between condition and age ($p = 0.052$; *SI Appendix*). Consistent with this interaction, our secondary preregistered analyses also found that 4-y-olds showed the predicted pattern (percent choice for drawing: 62.5% [Worse] vs. 18.75% [Better], $p = 0.029$, Fisher's exact), whereas 3-y-olds did not (25% [Worse] vs. 33.3% [Better], $p = 0.696$, Fisher's exact).

Given that age-related differences were not clear in prior experiments, what might explain these results? Our original intent in designing Experiment 4 was to create contexts in which the Observer had either a deflated (Worse Condition) or an inflated (Better Condition) representation of the child's performance. To appropriately manipulate this, the child's drawing must be clearly better than the bad drawing in the Worse Condition (making the Observer's belief deflated) and worse than the good drawing in the Better Condition (making the Observer's belief inflated). Rather unexpectedly, however, children's yoked evaluations of others' drawings suggested that many drawings did not meet these criteria; some drawings—especially those made by 3-y-olds—were rated by their peers to be even worse than the bad drawings (i.e., "the worst one"), suggesting that our key condition manipulation may not have had the intended effect for these children.

Although we did not list drawing quality as our preregistered exclusion criteria, as an exploratory analysis, we looked at children's choice of activity after excluding the following participants: those in the Worse Condition whose drawings were picked as the "the worst one" ($n = 13$, $M_{Age} = 3.89$; ten 3-y-olds; three 4-y-olds), those in the Better Condition whose drawings were evaluated as "the best one" ($n = 3$, $M_{Age} = 4.44$; one 3-y-old; two 4-y-olds), and those who did not receive another child's ratings on their drawings due to experimenter error or participants not responding to the question ($n = 5$, $M_{Age} = 3.91$; four 3-y-olds; one 4-y-old). We then analyzed the remaining group of children ($n = 43$ total; $n = 23$ in Better Condition, $M_{Age}(SD) = 4.16$; $n = 20$ in Worse Condition, $M_{Age}(SD) = 4.23$). Consistent with our prediction, we found that children were more likely to choose to continue drawing in the Worse Condition than in the Better Condition (55% vs. 21.7%, $p = 0.032$, Fisher's exact). As in Experiments 1 through 3, no effect of age was observed after the exclusion (*SI Appendix*).

In sum, Experiment 4 results suggest that preschool-aged children were sensitive to the desirability or valence of the Observer's self-relevant beliefs. Given that the Observer always had a false belief regardless of condition, children could have chosen to draw and reveal their true competence in both conditions. Instead, children were more likely to do so when such information could change the Observer's belief for the better. Note that these results can be interpreted in two different ways. Children might have had a novelty preference for blocks, which was attenuated specifically in the Worse condition by children's desire to reveal their actual drawing skills. Alternatively, those in the Better condition might have actively tried to maintain the Observer's inflated representation by purposefully withholding information about their drawing skills. These are not mutually exclusive possibilities, and more work is needed to further distinguish children's attempts to improve deflated representations versus maintain inflated representations. Unlike in previous experiments, we observed some effect of age in Experiment 4, but we caution against interpreting the age effect as reflecting genuine developmental change; age might be related to children's drawing abilities (and therefore the quality of their drawings), influencing the effectiveness of our condition manipulation. More specifically, children who made very low-quality drawings (who also tended to be younger) might have been discouraged from continuing to draw in both conditions, regardless of what the Observer thinks.

General Discussion

As humans, we are highly concerned with others' impressions of us and even try to change what others think of us. Bridging prior work on reputation management, theory of mind, and communication in early childhood, our results show that even young children can draw inferences about what others think of them and strategically communicate to manage others' representations of the self. Across four experiments, 3- and 4-y-old participants were able to track and use others' observations of their performance (e.g., their failures or successes, or the quality of their work) to

[¶]We did not ask participants to explain their choice in this experiment.

 https://doi.org/10.1073/pnas.2105642119

effectively revise others' outdated representations of the self (Experiments 1 through 3), selectively, when they could improve these representations to be positive (Experiment 4).[#] Notably, children made such decisions in the absence of any explicit mention or reminder of what others know or think about the self.

Despite much prior work on the presence of reputational concerns in children as young as age 3 (12, 13), the extent to which such findings reflect the ability to reason about others' minds has remained an open question. The current findings reveal the sophistication of children's reputational reasoning and the representations underlying their self-presentational behaviors. Children in our study did not rely on simple heuristics (e.g., behave positively in front of others), a learned association between positive behaviors and external reward (e.g., behave positively to receive praise), or a tendency to be consistent with an explicit reputation (e.g., acting "good" when they are told to have a reputation for being good); instead, they readily tracked others' observations of their performance and modulated their communication depending on how additional information would improve what others thought of them.

These findings connect prior work on reputation management (5) with existing developmental work on theory of mind (18) and informative communication (1). In particular, children's decisions about which toy to show to an observer (Experiments 2 and 3) reveal how children's mental-state reasoning capacities can give rise to different (and sometimes competing) communicative goals; without the ability to reason about others' mental states, children would not have understood the need to communicate information about the world (i.e., demonstrating how the toy works), nor the need to communicate information about the self (i.e., demonstrating their success). From this perspective, our findings fill important gaps in our understanding of theory of mind development and early communication.

First, prior work on theory of mind development has almost exclusively focused on children's reasoning about what others think about concrete, physical states of the world [i.e., where does Sally think her ball is? (17–22, 31, 32)]. Going beyond mental states about the physical world, our findings suggest that children can also reason about others' mental states about the self (i.e., what does Sally think about me?). Second, prior work on children as helpful communicators (1) has largely examined how children reason about others' mental states and utilities to provide information about the external world (i.e., telling Sally where her ball really is). Remarkably, children in our study capitalized on these representational and inferential capacities to communicate about the inner world (i.e., telling Sally that I actually can make the toy go). For adults, one of the most frequent topics of conversation is the self (41), and such self-disclosure may be particularly rewarding (42); our results suggest that young children are also motivated to communicate self-relevant information, in ways that reflect their ability to reason about others' minds.

These findings provide empirical evidence for attribution of self-relevant mental states in children as young as age 3. These results have broad implications about the nature of the representations that underlie early self-presentational behaviors. In Experiments 1 and 2, children only had to reason about whether the Observer(s) saw their success or not (ignorance representation).[‖] In Experiment 3, however, to selectively choose the Observed Toy over the Unobserved Toy, children had to attribute a representation that contains information about the specific toy that the child previously failed on. This could involve either a full-blown representation of belief (e.g., "She believes I cannot make this toy go") or a sophisticated form of propositional knowledge that contains information about a specific toy and the child's performance [e.g., "She doesn't know I made this toy go" (43)].[**] In Experiment 4, children had to go beyond representing what others know about their past performance and reason about how others might evaluate them (i.e., "good" or "bad" at drawing), suggesting that children had to attribute to the Observer a valenced representation of their drawing ability, or possibly a false belief. From this perspective, it is noteworthy that Experiments 3 and 4, which required attribution of contentful representations, produced relatively weaker results in this age group than Experiments 1 and 2, with suggestive age-related differences (see *SI Appendix*: *Effect of Age* for Experiments 3 and 4). Critically however, the current work does not provide conclusive evidence about the exact nature of representations children attributed to the observer across different experiments. These experiments also differed in other cognitive demands beyond mental-state attribution per se (e.g., number of toys, number of observers), and we did not find a clear relationship between children's performance in our task (Experiment 3b) and standard theory of mind measures (*SI Appendix*). Despite these limitations, these initial studies offer exciting directions for future work investigating the nature of the representations that underlie self-presentational behaviors.

Perhaps surprisingly, we did not find clear age-related differences in our experiments. Nonetheless, we can speculate on a few key factors that may underlie the developmental time course of children's ability to reason about what others think about the self. On one hand, representing others' beliefs about the self may be cognitively challenging (and correspondingly late developing), at least compared to representing others' beliefs about physical states of the world. Others' beliefs about qualities of the self, such as one's competence, morality, or preferences, are often abstract, subjective, and have no clear ground truth (46). Even understanding what kinds of observations (e.g., drawing a tree vs. lifting a box) can influence others' evaluations of different qualities of the self (e.g., artistic skills vs. physical strength) may require significant world knowledge and experience in specific domains. Thus, the ability to appreciate how the same observation (e.g., one's drawing) could be evaluated differently depending on the observers' standards or prior experiences may develop throughout childhood. Furthermore, although the current study used an experimental approach to directly manipulate the data children could use to infer others' representations of the self (i.e., an adult's observations of the child's failure or success), such data in a child's everyday environment may be much noisier. Others' observations of the self might be distributed across multiple people and extend over varying timescales, and, in some cases, may even require retrospective belief attribution (47) or counterfactual reasoning (48).

On the other hand, compared to others' knowledge or beliefs about the world, children may be particularly motivated to reason about self-relevant representations, especially when there are real-world consequences for their reputations (although, see ref. 49, which suggests that real-world consequences are not necessary). Arguably, the most compelling need for representing

[#]Findings from Experiments 3 and 4 had small effect sizes and should be interpreted with caution. Experiment 3b results were trending but not significant; Experiment 4 results were significant only in 4-y-olds, or after applying post hoc exclusion criteria.

[‖]This ignorance representation in Experiment 2 could lead to a preference for the Unobserved Toy in the Present Condition and equal preference between the two toys in the Absent Condition.

[**]Such knowledge-based representations might also reflect a sophisticated form of awareness relations (44) or experiential records (45).

others' mental states may arise in self-relevant contexts; to the extent that children care more about what Sally thinks about the self than where Sally thinks her ball is, we might find successful attribution of self-relevant mental states even before children succeed in standard theory of mind tasks. In fact, preschool-aged children interpret the same praise (e.g., "This drawing is great!") differently depending on the speaker's standards for praising other drawings (50), suggesting that children this age may already have a rather nuanced understanding of social evaluation in self-relevant contexts. Finally, although our task featured an observer whom children had not met before, children's motivation to consider what others think of the self may also be modulated by their relationship with the observer (e.g., parents, teachers, peers), how mistakes or failures are perceived by the self or others [e.g., mindsets (51)], and cultural values placed on various traits such as intelligence, generosity, or modesty [or an individual observer's values (52)].

Our findings raise questions for disciplines beyond developmental cognitive science. First, recent advances in computational models of social cognition have focused on formalizing how people infer what others think about the external world (53, 54); our work calls for computational work that formalizes people's inferences about representations of the inner world, and, more broadly, a computational framework for characterizing the process by which children build a coherent self-concept through their own experiences and social interactions. Second, in light of neural development of brain regions implicated in theory of mind (55, 56), exploring the role of these regions in representing others' mental states about the self (57) is another exciting direction for future work. Third, our findings shed light on classic findings on stereotype threat (e.g., ref. 3) as a byproduct of our ability to represent what others think of us, opening up opportunities for productive collaboration between developmental and social psychology. Finally, given recent advances in our understanding of social cognitive abilities in nonhuman primates (44, 58, 59) and domesticated dogs (60), one might ask whether analogous self-presentational behaviors are found in nonhuman species. It is possible that some species are sensitive to conspecifics' (or other humans') observations of the self and engage in self-promoting behaviors by demonstrating aggression or strength, although the extent to which these behaviors are modulated by rich representations of mental states remains to be seen.

The process by which we come to understand who we are may be intimately intertwined with others' understandings of who we are (6, 7). Just as we learn about the self from our own experiences, we also learn about who we are from interactions with others, and, in particular, what (we think) others think about us. The ability to reason about others' representation of the self may be especially critical in early childhood, allowing children to build the foundations for healthy self-concepts and social relationships with others. Our work is an initial step toward understanding how early inferential and representational abilities may play an important role in early learning about the self.

Materials and Methods

All study procedures were approved by Stanford University's Institutional Review Board. Parents gave informed consent for their children to participate. All studies took place in a private room in a university-affiliated nursery school.

Experiment 1.
Participants. Seventy 3- and 4-y-olds ($M_{Age}(SD) = 4.08(0.58)$, range: 3.00 to 4.99; 38 girls, 32 boys) were randomly assigned to the Think ($n = 35$) or Tell ($n = 35$) Condition. An additional 15 children were tested but excluded due to failing the memory check question ($n = 9$) or not completing the task ($n = 6$).
Stimuli. We constructed a novel cylinder-shaped causal device (a "toy") covered with blue felt, with two buttons on the sides. The experimenter used a hidden remote switch to manipulate children's apparent failures and successes. Children were also shown two 3″ × 5″ photos of the Failure Observer and Success Observer.

Experiment 2a.
Participants. Seventy 3-, 4- and 5-y-olds ($M_{age}(SD) = 4.66(0.66)$, range = 3.35 to 5.96; 39 girls, 31 boys) were randomly assigned to the Absent ($n = 38$) or Present Condition ($n = 32$). An additional 38 participants were tested but excluded for failing the memory check question ($n = 22$) or not completing the task ($n = 16$).
Stimuli. We constructed two novel causal toys, a red music toy and a green light-up toy (∼7″ × 7″ × 3″ each). Both toys had apparent causal mechanisms (red toy: two yellow buttons; green toy: two wooden levers) that generated music (red toy) or lights (green toy), but, in reality, the toys were controlled via a hidden remote switch. Children were also shown a 3″ × 5″ photo of the Observer.

Experiment 2b.
Participants. One hundred 3- and 4-y-olds ($M_{Age}(SD) = 4.00(0.53)$, range: 3.00 to 4.99; 57 girls, 43 boys) were randomly assigned to the Absent ($n = 52$) or Present Condition ($n = 48$). An additional 52 participants were tested but excluded for failing the memory check question ($n = 32$) or not completing the task ($n = 20$).
Stimuli. Toys used were similar to those in Experiment 2a, except that the music in the red toy was replaced with lights; thus, both toys had the same effect (lights) but different (fake) causal mechanisms.

Experiment 3a.
Participants. Thirty-two 3- and 4-y-olds ($M_{Age}(SD) = 4.01(0.59)$, range: 3.03 to 4.99; 17 girls, 15 boys) participated in a single condition experiment. An additional 15 participants were tested but excluded due to failing the memory check.
Stimuli. We used the same toys as in Experiment 2b.

Experiment 3b.
Participants. Fifty 3- and 4-y-olds ($M_{Age}(SD) = 4.03(0.59)$, range: 3.00 to 4.94; 28 girls, 22 boys) participated in a preregistered replication of Experiment 3a. An additional 12 participants were tested but excluded for failing the memory check.
Stimuli. We used the same toys as in Experiment 2b.

Experiment 4.
Participants. Sixty-four 3- and 4-y-olds ($M_{Age}(SD) = 4.12$, range: 3.43 to 4.97; 42 girls, 22 boys) were randomly assigned to the Better ($n = 30$) or Worse Condition ($n = 34$). An additional 19 participants were tested but excluded for failing the warm up question ($n = 6$) or memory check ($n = 13$).
Stimuli. "Good" and "bad" drawings were created using a blue marker on laminated sheets of paper. Children also drew with a blue marker on laminated paper. For the test question, children were shown a blue marker with a blank sheet of paper (drawing) and a clear, plastic tub with 10 (1″) wooden blocks (playing with blocks).

Data Availability. Anonymized children's responses data have been deposited in "Young children infer and manage what others think about them" (https://osf.io/uvmzy/) (61). All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank C. Dweck, M. C. Frank, S. Bridgers, N. Vélez, Y. Wu, A. Chuey, and members of the Social Learning Lab for helpful conversations and insightful comments about the project. We thank A. Garron, C. Norick, B. Anderson, G. Bennett-Pierre, J. Stegall, and B. Carrillo for help with data collection and coding. We are grateful to Bing Nursery School, to C. Yeh for help with recruiting and supervising testing, and to the children and families who participated. This work was funded by a James S. McDonnell Scholar Award and a Jacobs Foundation Early Career Research Fellowship (to H.G.) and an NSF Graduate Research Fellowship (to M.A.).

1. H. Gweon, *Inferential social learning: How humans learn from others and help others learn. Trends Cogn. Sci.* **25**, 896–910 (2021).
2. R. Rosenthal, L. Jacobson, Pygmalion in the classroom. *Urban Rev.* **3**, 16–20 (1968).
3. C. M. Steele, J. Aronson, Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* **69**, 797–811 (1995).
4. M. A. Nowak, K. Sigmund, Evolution of indirect reciprocity. *Nature* **437**, 1291–1298 (2005).
5. J. M. Engelmann, D. J. Rapp, The influence of reputational concerns on children's prosociality. *Curr. Opin. Psychol.* **20**, 92–95 (2018).
6. R. B. Felson, Reflected appraisal and the development of self. *Soc. Psychol. Q.* **48**, 71–78 (1985).
7. D. A. Cole, Change in self-perceived competence as a function of peer and teacher evaluation. *Dev. Psychol.* **27**, 682–688 (1991).
8. R. F. Baumeister, A self-presentational view of social phenomena. *Psychol. Bull.* **91**, 3–26 (1982).
9. M. R. Leary, R. M. Kowalski, Impression management: A literature review and two-component model. *Psychol. Bull.* **107**, 34–47 (1990).
10. I. M. Silver, A. Shaw, Pint-sized public relations: The development of reputation management. *Trends Cogn. Sci.* **22**, 277–279 (2018).
11. S. V. Botto, P. Rochat, Sensitivity to the evaluation of others emerges by 24 months. *Dev. Psychol.* **54**, 1723–1734 (2018).
12. L. Zhao, G. D. Heyman, L. Chen, K. Lee, Telling young children they have a reputation for being smart promotes cheating. *Dev. Sci.* **21**, e12585 (2018).
13. F. Ma, D. Zeng, F. Xu, B. J. Compton, G. D. Heyman, Delay of gratification as reputation management. *Psychol. Sci.* **31**, 1174–1182 (2020).
14. J. M. Engelmann, E. Herrmann, M. Tomasello, Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS One* **7**, e48433 (2012).
15. K. L. Leimgruber, A. Shaw, L. R. Santos, K. R. Olson, Young children are more generous when others are aware of their actions. *PLoS One* **7**, e48292 (2012).
16. G. Fu, G. D. Heyman, M. Qian, T. Guo, K. Lee, Young children with a positive reputation to maintain are less likely to cheat. *Dev. Sci.* **19**, 275–283 (2016).
17. H. Wimmer, J. Perner, Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103–128 (1983).
18. H. M. Wellman, D. Cross, J. Watson, Meta-analysis of theory-of-mind development: The truth about false belief. *Child Dev.* **72**, 655–684 (2001).
19. H. M. Wellman, D. Liu, Scaling of theory-of-mind tasks. *Child Dev.* **75**, 523–541 (2004).
20. K. H. Onishi, R. Baillargeon, Do 15-month-old infants understand false beliefs? *Science* **308**, 255–258 (2005).
21. P. Setoh, R. M. Scott, R. Baillargeon, Two-and-a-half-year-olds succeed at a traditional false-belief task with reduced processing demands. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13360–13365 (2016).
22. L. J. Powell, K. Hobbs, A. Bardis, S. Carey, R. Saxe, Replications of implicit theory of mind tasks with varying representational demands. *Cogn. Dev.* **46**, 40–50 (2018).
23. L. Kulke, M. Reiß, H. Krist, H. Rakoczy, How robust are anticipatory looking measures of theory of mind? Replication attempts across the life span. *Cogn. Dev.* **46**, 97–111 (2018).
24. S. T. Baker, A. M. Leslie, C. R. Gallistel, B. M. Hood, Bayesian change-point analysis reveals developmental change in a classic theory of mind task. *Cognit. Psychol.* **91**, 124–149 (2016).
25. C. Baer, O. Friedman, Fitting the message to the listener: Children selectively mention general and specific facts. *Child Dev.* **89**, 461–475 (2018).
26. I. Bass *et al.*, Children's developing theory of mind and pedagogical evidence selection. *Dev. Psychol.* **55**, 286–302 (2019).
27. H. Gweon, P. Shafto, L. Schulz, Development of children's sensitivity to overinformativeness in learning and teaching. *Dev. Psychol.* **54**, 2113–2125 (2018).
28. H. Gweon, L. Schulz, From exploration to instruction: Children learn from exploration and tailor their demonstrations to observers' goals and competence. *Child Dev.* **90**, e148–e164 (2019).
29. S. Bridgers, J. Jara-Ettinger, H. Gweon, Young children consider the expected utility of others' learning to decide what to teach. *Nat. Hum. Behav.* **4**, 144–152 (2020).
30. S. Ronfard, A. M. Was, P. L. Harris, Children teach methods they could not discover for themselves. *J. Exp. Child Psychol.* **142**, 107–117 (2016).
31. A. Gopnik, J. W. Astington, Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Dev.* **59**, 26–37 (1988).
32. C. M. Atance, D. K. O'Neill, Acting and planning on the basis of a false belief: Its effects on 3-year-old children's reasoning about their own false beliefs. *Dev. Psychol.* **40**, 953–964 (2004).
33. A. S. Smith-Flores, L. Feigenson, Preschoolers represent others' false beliefs about emotions. *Cogn. Dev.* **59**, 101081 (2021).
34. J. M. Engelmann, H. Over, E. Herrmann, M. Tomasello, Young children care more about their reputation with ingroup members and potential reciprocators. *Dev. Sci.* **16**, 952–958 (2013).
35. D. Stipek, S. Recchia, S. McClintic, M. Lewis, Self-evaluation in young children. *Monogr. Soc. Res. Child Dev.* **57**, 1–98 (1992).
36. H. Gweon, L. Schulz, 16-month-olds rationally infer causes of failed actions. *Science* **332**, 1524 (2011).
37. A. Cimpian, M. D. Hammond, G. Mazza, G. Corry, Young children's self-concepts include representations of abstract traits and the global self. *Child Dev.* **88**, 1786–1798 (2017).
38. K. Bartsch, H. M. Wellman, *Children Talk About the Mind* (Oxford University Press, 1995).
39. R. Magid, L. Schulz, Quit while you're ahead: Preschoolers' persistence and willingness to accept challenges are affected by social comparison. in *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, eds. D Noelle, *et al.* (Cognitive Science Society, Austin, TX, 2015), pp. 1482–1487.
40. A. Cimpian, H. M. C. Arce, E. M. Markman, C. S. Dweck, Subtle linguistic cues affect children's motivation. *Psychol. Sci.* **18**, 314–316 (2007).
41. R. I. Dunbar, A. Marriott, N. D. Duncan, Human conversational behavior. *Hum. Nat.* **8**, 231–246 (1997).
42. D. I. Tamir, J. P. Mitchell, Disclosing information about the self is intrinsically rewarding. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8038–8043 (2012).
43. M. Tomasello, How children come to understand false beliefs: A shared intentionality account. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8491–8498 (2018).
44. A. Martin, L. R. Santos, What cognitive representations support primate theory of mind? *Trends Cogn. Sci.* **20**, 375–382 (2016).
45. J. Perner, J. Roessler, From infants' to children's appreciation of belief. *Trends Cogn. Sci.* **16**, 519–525 (2012).
46. M. Asaba, A. Chuey, H. Gweon, Beyond knowledge versus belief: The contents of mental-state representations and their underlying computations. *Behav. Brain Sci.* **44**, e141 (2021).
47. I. Király, K. Oláh, G. Csibra, Á. M. Kovács, Retrospective attribution of false beliefs in 3-year-old children. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11477–11482 (2018).
48. A. Nyhout, P. A. Ganea, The development of the counterfactual imagination. *Child Dev. Perspect.* **13**, 254–259 (2019).
49. M. Asaba, X. Li, W. Q. Yow, H. Gweon, Children selectively demonstrate their competence to a puppet when others depict it as an agent. *Cogn. Dev.* **62**, 101186 (2022).
50. M. Asaba *et al.*, Young children use statistical evidence to infer the informativeness of praise. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, C. Kalish *et al.*, Eds. (Cognitive Science Society, Austin, TX, 2018).
51. C. S. Dweck, *Mindset: The New Psychology of Success* (Random House Digital, Inc., 2008).
52. G. D. Heyman, A. M. Compton, J. Amemiya, S. Ahn, S. Shao, Children's reputation management: Learning to identify what is socially valued and acting upon it. *Curr. Dir. Psychol. Sci.* **30**, 315–320 (2021).
53. C. L. Baker, J. Jara-Ettinger, R. Saxe, J. B. Tenenbaum, Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nat. Hum. Behav.* **1**, 0064 (2017).
54. J. Jara-Ettinger, L. E. Schulz, J. B. Tenenbaum, The Naïve Utility Calculus as a unified, quantitative framework for action understanding. *Cognit. Psychol.* **123**, 101334 (2020).
55. H. Gweon, D. Dodell-Feder, M. Bedny, R. Saxe, Theory of mind performance in children correlates with functional specialization of a brain region for thinking about thoughts. *Child Dev.* **83**, 1853–1868 (2012).
56. H. Richardson, G. Lisandrelli, A. Riobueno-Naylor, R. Saxe, Development of the social brain from age three to twelve years. *Nat. Commun.* **9**, 1027 (2018).
57. M. Kim, B. Park, L. Young, The psychology of motivated versus rational impression updating. *Trends Cogn. Sci.* **24**, 101–111 (2020).
58. C. Krupenye, F. Kano, S. Hirata, J. Call, M. Tomasello, Great apes anticipate that other individuals will act according to false beliefs. *Science* **354**, 110–114 (2016).
59. D. J. Horschler, L. R. Santos, E. L. MacLean, Do non-human primates really represent others' ignorance? A test of the awareness relations hypothesis. *Cognition* **190**, 72–80 (2019).
60. J. Kaminski, P. Piotti, Current trends in dog-human communication: Do dogs inform? *Curr. Dir. Psychol. Sci.* **25**, 322–326 (2016).
61. M. Asaba, H. Gweon, "Young children infer and manage what others think about them." Open Science Framework. https://osf.io/uvmzy/. Deposited 26 April 2017.