Who should I tell?
Young children correct and maintain others’ beliefs about the self

Mika Asaba (masaba@stanford.edu)  Hyowon Gweon (hyo@stanford.edu)
Department of Psychology, Stanford University, Stanford, CA 94305

Abstract
We care tremendously about what other people think of us. Motivated by two lines of prior work – children’s inferential and communicative capacities and strategic reputation management – we examine how children infer what others think of them given others’ observations of their performance, and how they influence these beliefs through discerning their performance. In Experiment 1, 3-5 year-olds played a luck-based game; one confederates watched the child win and another confederate watched the child lose. We asked the child to disclose an additional, unobserved win to one of the two confederates. We find that younger children overwhelmingly choose the person who previously saw them win. However, as age increased, children were more likely to choose to disclose to someone who previously saw them lose. In Experiment 2, adults played a similar third person version and selectively chose the person who saw the main character previously lose.

Keywords: Theory of Mind; social cognition; cognitive development; communication; reputation management

Introduction
We are deeply curious about the minds of other people. From merely observing others’ behaviors to probing them with direct questions, we frequently engage in learning about others’ beliefs, and even attempt to change these beliefs by communicating with others. Among many kinds of unobservable contents of other people’s minds, there is one suite of beliefs that we care extraordinarily about: others’ beliefs about us.

Decades of research on Theory of Mind (ToM) – our capacity to understand and reason about others’ unobservable mental state – has revealed much about both its early emergence (Baillargeon, Scott, & He, 2010) and its developmental trajectory (Wellman, Cross, & Watson, 2001). A host of prior work has examined various domains in which ToM plays a central role in our everyday social interactions, such as helping others (Buetellmann, Carpenter, & Tomasello, 2009), morally evaluating others (Cushman, Sheketoff, Wharton, & Carey, 2013), and teaching and communicating with others (Gweon, Pelton, Konopka, & Schulz, 2014; Strauss & Ziv, 2012). In this study, we highlight another important role of ToM in our social lives: it allows us to reason about others’ beliefs and evaluations of us, and help us change or maintain these beliefs by informing others about ourselves.

Imagine the following scenario: your friend Tom observes you win many amazing tennis matches, but later your friend Sam watches you lose several games in a row. Based on what they have seen, you might reasonably infer that your friends would hold different beliefs about your skill level; Tom would think of you as a competent tennis player, while Sam might think quite the opposite. Based on these inferences, you may even attempt to correct Sam’s belief (e.g., by telling him that you recently won many matches) while maintaining Tom’s belief (e.g., by omitting your losses). We propose that these seemingly simple intuitions emerge from a sophisticated ability: (1) infer what others believe about us given some evidence about the self, (2) understand that such beliefs can be revised and updated given new information, and (3) decide how to selectively communicate information about the self to cultivate or maintain others’ positive beliefs about us.

Previous developmental work has provided evidence for young children’s competence in each of these components outside of the domain of beliefs about the self. First, children draw inferences about others’ beliefs, epistemic states, and even moral dispositions based on their experiences and behaviors (Hamlin, 2013; Koenig & Harris, 2007). Second, children understand that others’ beliefs can change with additional data (Song, Onishi, Baillargeon, & Fisher, 2008), and that these revisions are sensitive to the amount and strength of evidence (Schulz, 2012; Gweon, Shafto, & Schulz, 2014). Third, preschool-aged children also understand that they can actively cause such revisions to others’ beliefs by selecting and communicating information for others (Ding, Wellman, Wang, Fu, & Lee, 2015; Gweon, Chu, & Schulz, 2014; Gweon, Shafto, & Schulz, 2014; Rhodes, Bonawitz, Shafto, Chen, & Caglar, 2015). For instance, children differentially select evidence to show others, depending on whether the goal is to teach or deceive (Rhodes et al., 2015) and the learner’s prior beliefs (Gweon, Shafto, & Schulz, 2014).

However, in most of these studies, the contents of others’ beliefs involve states of the external world with a ground truth (e.g., locations or names of objects, causal mechanisms). For instance, in the Sally-Anne task or other variations of false-belief tasks (Wimmer & Perner, 1983; Gopnik & Astington, 1988), children are questioned about the contents of others’ beliefs that are either accurate or inaccurate with respect to the current states of the world. However, in people’s beliefs about the self (or people’s beliefs about people more generally), there is not always a clear ground truth; these beliefs are formed via various sources of information that reflect a person’s unobservable, internal qualities (e.g., competence, niceness, knowledgability, etc.). Therefore, the accuracy of such beliefs can only be assessed with respect to the representativeness of such observations. Therefore, it remains an open question whether children’s competence in reasoning about others’ factual beliefs about the world also naturally extend to reasoning about others’ beliefs about the self.

In reasoning about others’ beliefs about us, we not only care about the accuracy of others’ beliefs but also about the desirability of the beliefs; that is, we care tremendously about whether others think we are competent, nice, and knowl-
edgeable, independent of its ground truth. Indeed, prior research on children’s impression management suggests that even young children are motivated to create a positive reputation with others (Shaw, Li, & Olson, 2013; Lee, 2013), using appropriate communicative (e.g., lying) and behavioral (e.g., sharing, helping) means to do so. For instance, even 3-year-old children lie about peeking into a box when an experimenter told them not to do so (Chandler, Fritz, & Hala, 1989; Polak & Harris, 1999). Five-year-olds help more and cheat less in the presence of others (Engelmann, Herrmann, & Tomasello, 2012; Leimgruber, Shaw, Santos, & Olson, 2012), and even when they have been told that other children think they are good (Fu, Heyman, Qian, Guo, & Lee, 2015). By the end of the preschool years, children show both a motivation for others to think well of them as well as a sensitivity to particular kinds of information that might lead others to form positive impressions of them in prosocial or moral contexts.

While these studies have demonstrated children’s desire to maintain a positive reputation with others, most of these studies looked at children’s behaviors with others with whom they had no prior experience; that is, previous work has mainly shown children’s attempts to create a good “first impression”. However, many of our daily interactions involve people we have already interacted with before, and thus have already formed some beliefs or impressions of us. If we want these individuals to hold positive beliefs about ourselves, we must employ our Theory of Mind to manage these beliefs: we need to infer what others believe about us, and when such beliefs are discrepant from what we want others to believe, we must provide additional information to change these beliefs.

As adults, we intuitively understand that one can be selective about to whom we tell different pieces of information. Importantly, such selective and targeted communication of information is driven by our desire for others to know only certain (and often desirable) aspects of ourselves. We also recognize that the same piece of information (e.g., winning a game) can differentially impact what others think of us. In the tennis example, disclosing to Sam (who thinks you are not very skilled) that you won a game might greatly improve his beliefs about your skills, whereas telling Tom (who thinks you are good) might simply reaffirm what he thinks of you.

In our study, we asked whether young children prefer to provide new positive information about themselves to someone who holds a less positive belief about their performance. Older children might understand the value of disclosing a win to someone who already holds a positive belief about them, whereas older children might understand the value of disclosing a win to someone who holds a less positive belief about their performance.

### Method

**Subjects** Seventy-three preschool-aged children ($M_{Age} (SD)$: 4.56 (.71), range: 3.23 - 5.98) were recruited from an on-campus preschool. An additional eleven children were tested and dropped from analysis due to failure to complete the study ($N = 7$) or experimenter error ($N = 4$).

**Materials** Children played a simple game with the experimenter using laminated cards (2”x 3”). The winning cards were red on the front with a large yellow star in the middle, while the losing cards were blue on the front with no star; all cards were yellow on the back. Two headshot photos (3” x 5”) of each confederate wearing neutral colored clothing. Stickers were used as rewards. See Figure 1B for stimuli.

**Procedure** Participants were tested in a quiet room inside of the preschool. First, the experimenter introduced the game to the participant: in each trial, the experimenter would lay out several cards face down on the table, and the child could choose one of the cards to flip over. Then, the experimenter showed the two types of possible cards: a red card with a star on it meant that the child won (i.e., win a sticker); a blue card with no star meant the child lost (i.e., receive nothing). The experimenter asked the child what happens if each type of card was picked; all children were able to correctly report the outcome of each card. The experimenter showed the child the two pictures of the confederates, provided their names (e.g., Anne and Sally), and said, “Sometimes, my friends will come in and watch the game. Does that sound okay?” All children agreed for the confederates to come in during the game.

Children played four trials of the game; the cards were stacked such that the experimenter could control the outcome for each trial. If the child chose a winning card, the experimenter said, “Oh wow, you got a star! That means you get a sticker!” If the child chose a losing card, the experimenter said, “Oh no, that means you do not get anything.” The child always won on the first trial. Before each of the second and third trials, the experimenter told the child, “Okay, I think my friend Anne (Sally) wants to watch now!” And then she
looked at the door (behind the child) and said, “Hey Anne (Sally), is that you?” Anne (Sally) then entered the room, smiled politely, and sat in a chair between the child and experimenter (see Figure 1 for schematics of the room setup). The confederate maintained a neutral to mildly positive facial expression, made no comments while the game was playing, and did not respond to the outcome of the game. The child won in front of one confederate and lost in front of another (order and confederate counterbalanced). After the outcome was revealed, the confederate left the room.

On the final, fourth trial, no confederate was present and the child always won. Then, the experimenter brought out the pictures of the confederates and told the child that now he or she can tell Anne or Sally about the final win. The child indicated his or her response by saying a name or pointing to one of the photos; if the child did not respond or said both, the experimenter asked the child to only choose one confederate. The experimenter called the chosen confederate into the room. After the child told the confederate what happened, the confederate left the room and the child then received three stickers for the three winning trials.

Results and Discussion

Three- to five-year-olds were asked to disclose the final win to either someone who previously saw them win (“Win” observer) or someone who previously saw them lose (“Loss” observer). Our main interest was which confederate children chose, and whether children’s choices change with age. We found no effects of gender, the order of the win and lose in the second and third trials, or the identity of the confederates. Collapsing across all participants, we found that 31 of 73 participants (42.5%) chose to disclose the final win to the “Loss” observer ($p = .242$). Next, we used the median age (4.64) to split participants into a younger (N = 36; $M_{Age}(SD): 3.96(4.3))$ and older group (N = 37; $M_{Age}(SD): 5.14(3.3))$. We found that only a small proportion of the younger group, 11 of 36 (30.5%) chose to disclose the final win the “Loss” observer ($p = .029$, by binomial test). However, the older group showed a different pattern; although the number of children who chose the “Loss” observer was not significantly above chance ($p = .743$, by binomial test), they were significantly more likely to disclose to the “Loss” observer compared to the children in the younger group (20 of 37 or 54.1%, younger vs. older: $p = .059$, by fisher’s exact test). Indeed, in a logistic regression model ($1 = Chose “Loss”, 0 = Chose “Win”) Age was a significant predictor in children’s choice of confederates ($\beta = .828, p = .025$), with older children were more likely to disclose to the “Loss” observer.

In sum, we found that although younger preschool-aged children preferred to disclose the final win to the “Win” observer, children gradually become more likely to disclose to the “Loss” observer. Interestingly, the “Win” observer already had a positive belief about the child, and thus disclosure of an additional win would do little to change her beliefs. While older children were significantly more likely to disclose to the “Loss” observer, the proportion of children who did so was only slightly above chance. Moreover, this result shows young children’s striking ability to distinguish between the two observers after only single trials.

Given the results from the logistic regression, one question to further explore is when we might preferentially select the “Loss” observer over the “Win” observer. In the next study, we asked whether even adults would choose the “Loss” observer. It is also interesting to consider how the nature of the game (e.g., luck-based vs. skill-based games) might differentially affect motivation to correct others’ beliefs. Here, children played a simple, luck-based game but if the game’s outcome was determined by effort or skill, there might be a greater concern for reputation and thus a stronger tendency to maintain or change others’ beliefs in the desired direction. In Experiment 2, we first investigate adults’ choices in an online third-person version of the task (2A), and directly compare participant’s responses in a game of chance and a game that involves more skill (2B).

**Experiment 2A: Adults (Luck-Based)**

In Experiment 2A, we first asked about adults’ preferences to correct others’ negative impressions. In order to make the scenario plausible for an online study, we asked people to predict another agent’s choice to disclose, rather than their own choices. Adults observed an agent (Bill) succeed or lose in a card game task as two confederates observed the agent at different times. We asked the participants to choose who Bill should tell about his final successful performance: the friend
who watched him win, or the friend who watched him lose.

Methods
Subjects Adults (N=100) were recruited from Amazon’s Mechanical Turk (AMT), M_{Age}(SD): 34.2(8.9), range: 22 - 60. Fifteen additional participants were tested but excluded from analysis due to failing a planned attention-check question (see Procedure).

Materials Participants observed several cartoon pictures in the online game. The game consisted of a vertical board with five cards on it (see Figure 2); the dealer is depicted in green. A card with a star on it meant that Bill won; a card with an X on it meant that Bill lost. Only the characters backs were shown during the observation pictures.

Procedure
Participants first learned how the game worked and what each card meant; only those who correctly reported that the winning card has a star were included in analyses. We ran two versions of the study in which we varied the amount of evidence. In one version (N=50), Bill won once by himself, then won once in front of one friend (“Win” observer) and lost once in front of the other (“Loss” observer). Each win or loss trial was represented in its own image on a page. For the fourth and last trial, Bill won once all by himself. Then, participants were asked who Bill will tell about his final win: the friend who previously saw Bill win once or the friend who previously saw Bill lose once. Participants were asked to explain their choice afterwards.

In another version (N=50), Bill won or lost three trials in a row in front of each confederate, and participants were asked to choose whom Bill should disclose his final three wins. For both versions, we counterbalanced the order of the observed outcomes (W L vs. L W; W W W L L L vs. L L L W W W) as well as which of the two friends observed during these trials.

Results and Discussion
We found no effect of the order of the observed outcomes (W first vs. L first), the order of the friend (Anne or Sally), or the number of trials (1 or 3 trials). Thus the data were collapsed for further analysis. Participants were asked to choose which one of the two friends Bill should tell about his final win: the friend who previously saw him win lose before versus the friend who previously saw him lose before. 61% (61 of 100) of participants chose the person who previously saw Bill lose (binomial test, \( p = .03 \)). Thus adults were sensitive to the kind of outcome each friend observed, and predicted Bill to selectively disclose to the friend who previously saw him lose.

Experiment 2B: Adults (Skill-Based)
In Experiment 2B, we asked how manipulating the nature of the game (luck-based vs. skill-based) might affect adults’ reasoning about selective disclosure. When the game involves skill rather than luck and is thus more informative about the player’s underlying competence, more participants might predict that the agent would communicate his positive performance to someone who holds a negative belief about his competence than someone who holds a positive belief.

Methods
Subjects Participants (N=200) were recruited from AMT, M_{Age}(SD): 39.1(10.1), range: 24 - 59. Twelve participants were excluded from analysis due to failing the two planned attention-check questions (see Procedure).

Materials As in Experiment 2A, participants observed a cartoon scenario where Bill played a game as one of his friends (Anne or Sally) was present. The game featured a dart board divided into 6 alternating blue and red slices.

Procedure
The structure of the task was nearly identical to Experiment 2A, and the only difference was that Bill played darts, rather than a card game. This allowed us to easily manipulate the nature of the game (luck-based vs skill-based) using nearly identical stimuli. Participants were randomly assigned to the Skill and Random conditions (N = 100 in each condition). In the Random condition, Bill simply pressed a button to activate a dart-throwing machine that would place the dart at a random location on the board. In the Skill condition, Bill aimed at the board to throw the dart, but no information was provided about his competence. In all conditions, participants were told that Bill would “win” if the dart landed on one of the blue slices on the board, and “Loss” if it landed on red.

As in Experiment 2A, we ran two versions of this task. In one version, one of Bill’s friends observed him win once and the other friend watched him lose once. Then, Bill won once in the last weekend when no one was watching. We asked participants who Bill will tell, the friend who observed him win before or the friend who observed him lose before. In another version, Bill won or lost three times (instead of once) in front of each confederate and would disclose three wins to one of them. We counterbalanced the order of the outcomes and which of the two friends observed during these trials.

Results and Discussion
There was no effect of the order of the observed outcomes (W vs. L), the order of the friend (Anne or Sally), or the number of observation trials (one or three); thus the data were collapsed for further analysis. As in previous experiments, participants were asked to which friend Bill should tell about his win. In the Random condition, 68% (68 of 100) of participants chose the person who previously saw Bill lose (binomial test, \( p < 0.001 \)), replicating the results from Experiment 2A. In the Skill condition, 71% (71 of 100) of participants chose this person (\( p < 0.001 \), by binom. test). Contrary to our prediction, there was no difference between the two conditions.
General Discussion

In this study, we explored how people communicate information about the self to others. More specifically, we asked whether children as well as adults would choose to disclose positive information about the self (e.g., winning a game) to someone who previously observed a negative outcome or to someone who observed a positive outcome about the self. If people understand that: (1) others’ prior observations of the self can result in positive or negative beliefs about one’s competence, and (2) one can revise such beliefs by communicating information about the self, they would be sensitive to others’ prior observations and preferentially choose to communicate to the person who observed a negative outcome.

In Experiment 1, we tested our prediction with children between ages 3 - 5. We found that while younger children preferred to disclose their positive performance to someone who has previously seen them win rather than lose, older children were more likely to choose the person who has previously seen them lose. While older children’s choice was not significantly above chance, children’s age predicted their choice. These results suggest that even though young children might be biased to prefer someone who already holds a positive belief about them, children gradually understand the benefit of correcting others’ negative beliefs. Under this interpretation, the chance-level choice in older children might suggest that they were torn between their preference for one observer (i.e., “Win”) and their desire to communicate their winning outcome to the other observer (i.e., “Loss” observer).

One alternative explanation is that young children’s preference for someone who holds a positive belief about them gradually disappears; in this case, older children’s choice might reflect a genuine absence of preference between the two observers. However, intuitively even adults prefer people who hold positive impressions of them, and it is unlikely that such preference is absent in the older half of the children in our study. Future studies are needed to provide stronger support for the idea that our results reflect children’s developing understanding that others’ beliefs about the self can be revised and maintained via selective disclosure of information.

In Experiment 2, we tested adults’ intuitions in analogous tasks and further explored the role of luck vs. skill in these inferences. Adults showed a systematic preference for the person who observed a negative outcome of the game, suggesting that they understand that one can change others’ negative beliefs about the self by communicating positive information about the self. While we failed to find a difference between the luck-based and skill-based games, it is possible that the positive glow from a lucky win was just as powerful as a skill-based outcome for adults, or that the skill-based game also increased the desire to maintain people’s positive beliefs about an underlying competence. We also did not find a difference in selectivity for the “Lose” observer when we varied the amount of evidence (one vs. three instances), but this may have also been due to an increase in desire to affiliate with the person who previously viewed multiple wins.

Future studies will further explore how the nature of these games and amount of observed evidence might differentially affect others’ impressions of oneself. For instance, in the face of someone who holds a less positive belief about the child, we could ask whether the child would like to disclose a luck-based positive outcome or a skill-based positive outcome. Further, we could compare how children differentially disclose information to others who have weak positive (e.g., viewed one win) or strong positive (e.g., viewed multiple wins) evidence of the child’s performance.

Previous work in children’s understanding of disclosure have utilized relatively complicated vignettes and asked children to either endorse or reject others’ disclosure about another agent to a third party agent (Shaw & Olson, 2015) or predict whether or not others will disclose information about themselves (e.g., Hicks, Liu, & Heyman, 2015). Our novel and simple paradigm provides a rich and naturalistic environment from which we can better understand how children infer others’ beliefs of them based on what others have observed, and how they communicate information about themselves to change or maintain these beliefs.

Last, we hope that this work will inspire future research integrating work in social cognitive development and impression management behaviors. While reputation management behaviors have largely concerned valence-based impressions
(e.g., “good” or “bad”), there are many other kinds of beliefs that others can have about us. For instance, even our preferences or a set of beliefs can provide rich information about who we are, and selectively communicating these to others can be an effective way to manage the representations of ourselves that others hold.

Inferring others’ beliefs about the self and knowing how to change and maintain them might be particularly important for young children who are just beginning to develop a coherent self-representation and forming initial relationships with others. This capacity to reason about others’ beliefs of us is deeply important for navigating our daily communicative interactions, building meaningful, lasting bonds with others.

Acknowledgments

We thank Sumudu Rathnayake, Sara Altman, Chelsey Pan, and Amy Qian for help as confederates. We also thank Bing Nursery School and the children and families who participated. This work was supported by a Varieties of Understanding grant from the John Templeton Foundation to HG.

References


Shaw, A., Li, V., & Olson, K. (2013). Reputation is everything. *Navigating the social world*.


