Children expect adults to hold gender stereotypes, even when they are not accurate

Mika Asaba (mika.asaba@yale.edu)[1],
Marianna Zhang (marianna.zhang@stanford.edu)[2],
Julia Leonard (julia.leonard@yale.edu)[1],
Department of Psychology, Yale University, New Haven, CT 06511
Department of Psychology, Stanford University, Stanford, CA 94025

Abstract

Gender stereotypes are early-emerging and harmful for young children. However, it is unclear how children reason about other people's gender stereotypes, especially when they differ from children's own beliefs. Across two preregistered experiments (total n=271), we tested whether 5- to 7-year-old children expect teachers to give engineering games to boy students and story games to girl students, even when children themselves know that these are not students' true preferences. Experiment 1 found that participants were more likely to predict that a teacher would give students stereotypical games when the teacher did not know (versus did know) the students' true counter-stereotypical interests. In Experiment 2, when the students expressed interest in both games, 6- and 7-year-olds selectively predicted that teachers would give students whom they had just met stereotypical games. Thus, by the time children enter school, they think that adults hold gender stereotypes, even if children know these stereotypes are inaccurate, which may impact children's learning and decision-making in the classroom.

Keywords: Gender Stereotypes; Children; Social Cognition, Theory of Mind

Introduction

Stereotypes about social groups are early-emerging, pervasive, and harmful. By the early school years, children in the United States form stereotypes regarding gender (e.g., Cvencek, Meltzoff, & Greenwald, 2011), race (e.g., Pauker, Xu, Williams, & Biddle, 2016), and class (e.g., Sigelman, 2012). Importantly, children's stereotypes can inform what they think of their own abilities (Bian, Leslie, & Cimpian, 2017) and which activities or careers they decide to pursue (Master, Meltzoff, & Cheryan, 2021a). While much work has been devoted to investigating how children develop stereotypes, less is known about how children reason about *other people's* stereotypes.

Understanding how young children think about other people's stereotypes is valuable for informing theories about stereotype development and stereotype threat. First, awareness of others' stereotypes may strengthen children's own stereotypes or enable them to develop in the first place. Consistent with this, some prior work has revealed positive correlations between explicit awareness of other people's stereotypes and school-aged children's own stereotypes (e.g., Cvencek, Nasir, O'Connor, Wischnia, & Meltzoff, 2015). Thus, understanding whether and when young children expect others to hold stereotypes is critical to scientific theories about stereotype development. Second, reasoning about others' stereotypes may impact children's own behaviors and performance, especially in learning contexts. Prior work on stereotype threat has shown how stereotypes can become self-fulfilling prophecies when people become worried about confirming a negative stereotype and, in fact, perform worse due to this worry (e.g., Spencer, Steele, & Quinn, 1999; Spencer, Logel, & Davies, 2016). Thus, thinking about other people's stereotypes may cause children who feel the weight of those stereotypes to perform worse in the classroom, even if children themselves do not hold the stereotype. Here we ask, do young children expect adults to hold stereotypes, and what are the cognitive capacities that support these expectations?

Prior work has provided mixed results on children's awareness of adults' stereotypes. For example, Freeman (2003) suggested that 3- and 5-year-old children thought that their parents would be less approving of gender counter-stereotypical toy choices (e.g., a tea set for a boy) than stereotypical toy choices. Other studies have asked adolescents or school-aged children explicit questions about what "most adults" think of girls and boys (e.g., "I think that in math and science, most adults think boys are much better than girls", Kurtz-Costes, Copping, Rowley, & Kinlaw, 2014). Some work has shown that children's awareness of stereotypes increases during the school years, with stigmatized groups showing earlier knowledge (at least of racial stereotypes, McKown & Weinstein, 2003). However, other work has shown that fourth graders show an own-gender bias for others' stereotypes (e.g., girls think adults think girls are more competent), or think that adults hold egalitarian beliefs (Kurtz-Costes et al., 2014).

Beyond these mixed findings, a key question remains: How are children reasoning about other people's stereotypes? In prior work, children's own stereotypes were never differentiated from their expectations of others' stereotypes. Thus it is unclear whether children assume that others simply hold the same stereotypes that they do, or whether they are *representing* others' stereotypes as mental states distinct from their own. If the latter, then this would mean that even children who do not hold stereotypes could still expect others to hold them. This representation is important to uncover, as it may be a key source of how children form their own stereotypes and in turn, scaffold children's learning behaviors. To test whether children represent other people's stereotypes as separate from their own beliefs, it is critical to set up situations where children themselves do not hold stereotyped beliefs about a target agent (e.g., if they know that the agent

has counter-stereotypical interests) but other people may still have them.

Decades of research from the Theory of Mind literature have shown that young children can represent others' beliefs as separate from their own (e.g., Wellman, Cross, & Watson, 2001; Wellman & Banerjee, 1991; Gopnik & Astington, 1988). For example, 3-to-4-year-old children understand that other people can hold different beliefs about their competence, compared to their own beliefs, depending on others' observations of their performance (Asaba & Gweon, 2022). Furthermore, 6-to-8-year-old children understand that they shouldn't follow the advice from an adult who wrongly thinks they are highly competent at a certain task (Bass, Mahaffey, & Bonawitz, 2021). However, this literature has primarily focused on inferences about people's beliefs based on previously observed evidence (e.g., prior performance). Thus, it is not clear whether children expect people to hold assumptions or prior beliefs about new individuals based solely on their group membership (e.g., their gender) and not any observable evidence.

Here, we bridge work on children's stereotype awareness with research on children's mental state reasoning to ask whether young children expect adults to hold stereotypes, even when they know they are not accurate. Imagine the following: a child knows that there's a girl who loves engineering (a counter-stereotypical domain). When a teacher meets this girl for the first time, what would the child predict that the teacher thinks about this girl and what she likes? Would they predict that the teacher is more likely to think that she prefers stories and reading (stereotypical) over engineering (counter-stereotypical)? One possibility is that children do not expect adults to hold prior beliefs about novel individuals and thus they should be at chance on this question. However, given that children can represent others' beliefs (Wellman et al., 2001), are sensitive to adults' non-verbal (stereotyped) treatment of others, (Brey & Pauker, 2019), and are enculturated in a stereotyped world (Bigler & Liben, 2006), we predict that children will assume that adults hold prior beliefs about novel individuals based on their social group (e.g., gender), even in contexts where children know these stereotypes are not warranted (e.g., they know the student in question likes counter-stereotypical tasks). This would result in them guessing that the teacher would give the girl story games.

We specifically focus on gender stereotypes about interest in STEM (i.e., that boys are more interested in engineering than girls) given that such stereotypes are early emerging and predictive of children's task choices (Master, Meltzoff, & Cheryan, 2021b). We also focus on children ages 5-to-7, because (i) children this age are capable of representing others' mental states (Wellman & Liu, 2004) and (ii) prior work has shown that by around age 6, children hold gender stereotypes about interest (Master et al., 2021b).

In Experiment 1, we assessed whether 5- to 7-year-old children expect teachers to give "engineering games" to boy students[1] (given that engineering tends to be stereotyped in favor of boys; Master et al., 2021b) and "story games" to girl students (given that reading tends to be stereotyped in favor of girls; Retelsdorf, Schwartz, & Asbrock, 2015). Critically, while participants themselves always knew whether a particular student was more interested in engineering or stories, we manipulated whether the teacher knew this information, thus allowing us to differentiate the teacher's beliefs from participants' own beliefs. Experiments 2a (6-year-olds) and 2b (7-year-olds) were largely similar, with some methodological changes to constrain our interpretations from Experiment 1. For all experiments, we included a measure of children's own stereotypes to explore the relationship between their own stereotypes and their reasoning about other people's stereotypes. All experiments were preregistered, and preregistrations, data, and analyses can be found here: http://tinyurl.com/stereoRepresentation.

Experiment 1

Methods

Participants We recruited 145 5- to 7-year-old participants in the United States (mean age: 6.30 years, *SD*: 0.83, range: 5.01-7.97) on *Lookit*, an asynchronous online testing platform (Scott & Schulz, 2017). Parents reported participants' gender (51.7% boys, 48.3% girls) and race (55.9% White, 17.2% Asian, 7.6% Asian and White, 2.8% Hispanic, Latino, or Spanish origin, 2.1% White and Hispanic, 2.1% Black, 2.1% White and Native Hawaiian or Other Pacific Islander, and 10.2% other categories). An additional 5 participants were recruited but excluded for living outside of the US, and all other participants passed our exclusion criteria.

Stimuli In the test trials, four images of adults (two women, two men) and four images of children (normed on age and emotional expression) were used as the teachers and students. These images were overlaid on a cartoon classroom that showed a door and two computer games (one with an engineering icon, and one with a story icon); see Figure 1. In the stereotype trials, simple line drawings of a boy and a girl were presented, and blue thumbs up and red thumbs down icons were shown as response options.

Procedure Children participated in the study remotely with their parents or legal guardians on *Lookit*, with no experimenter present. Participants first underwent a brief tutorial. They were introduced to "Sunny School", where it is the first day of school, and thus, teachers have not met their students before. Then, participants were introduced to the two types of games in the classrooms: story games (where "students make up or act out stories") and engineering games (where "students create or build things, like machines or buildings"), each with their own icon depicting the game (see Figure 1).

[1] Given that work on gender interest stereotypes has focused on girls and boys, we present boy and girl characters in our studies. Thus, our work is limited in its characterization of stereotypes, or expectations of stereotypes, about non-binary children.

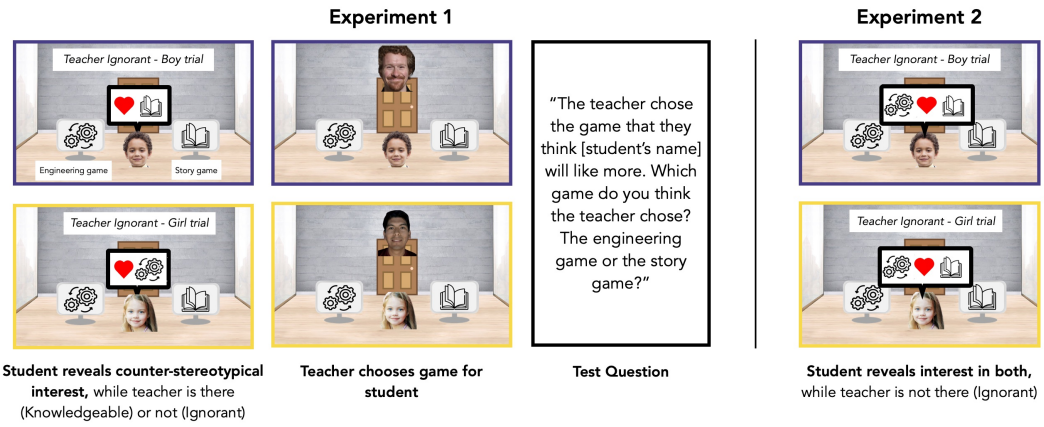

Figure 1: Stimuli from Exp. 1-2. In Exp. 1, a student (a girl or boy) revealed their counter-stereotypical interest (i.e., boys like stories, girls like engineering). Critically, their teacher was either there (Teacher Knowledgeable) or not (Teacher Ignorant) when the student said their interest. Then, participants predicted which game the teacher chose for the student based on what they think the student likes. Participants each underwent four trials that crossed the gender of the student (girl or boy) and the teacher's knowledge of their interest (Knowledgeable or Ignorant). Exp. 2 was identical, except the students said that they liked *both* the engineering and story games, and the teacher was always ignorant.

To make sure participants were listening, they were asked to say out loud what happens in engineering and story games.

Participants were asked three check questions before continuing to the test trials. First, they were asked to click on the correct icon for each game (response options were the engineering icon and the story icon; 2 questions). Then they were asked whether the teachers had seen the students play before (response options were a thumbs up or a thumbs down icon; 1 question). For all questions, participants received feedback (e.g., "That's right, the teachers haven't seen them play before" or "Actually, the teachers have seen them play before"). Participants were required to answer all three check questions correctly before continuing the task and were given unlimited chances. We preregistered excluding any children who failed to correctly respond to the game icon questions after one correction; no participants were excluded for this.

Next, participants underwent 4 test trials that varied the teacher's knowledge of the student's interest (Teacher Knowledgeable or Teacher Ignorant) and the gender of the student (boy or girl). In each trial, participants met a boy or girl student who stated their interest in engineering or stories (e.g., "Here is Nathan. Nathan is looking at the engineering game and the story game. Nathan says 'I really like stories!'"). As a check question, participants were asked to click the game that the student likes (all participants passed these check questions). In the Teacher Knowledgeable trials, the student's teacher then entered the room and the student stated their interest (such that the teacher heard the student's interest). In the Teacher Ignorant trials, the teacher entered *after* the student stated their interest (such that the teacher did not hear it). At test, participants were asked to predict which game the teacher chose for the student: e.g., "The teacher chose the game that he thinks Nathan will like more. Which game do you think the teacher chose for Nathan?" Participants responded by clicking on the engineering or story icon.

Notably, the students always expressed interest in the counter-stereotypical game—that is, the boys were always interested in stories and the girls were always interested in engineering. This design choice allowed us to test whether participants were genuinely relying on their expectations for the teacher's beliefs about the student's interests to make the game prediction, rather than what the student was actually interested in. Trials were blocked by the knowledge manipulation (order counterbalanced; Teacher Ignorant trials first or Teacher Knowledgeable trials first), with student gender order (boy first or girl first) and teacher gender (e.g., male teacher with male student, or male teacher with female student) counterbalanced within the blocks.

If children do expect adults to hold gender stereotypes, then we should see an interaction between the teacher's knowledge and the student's gender in participants' predictions of what game a teacher chose for a student. Specifically, we predicted that, in the Teacher Knowledgeable trials, participants would reliably expect that the teacher would select the engineering game for the girl and the story game for the boy, given that the teacher knows that these are what each student is interested in. In the Teacher Ignorant trials, however, there are a few possible patterns of evidence. The strongest evidence would be if this pattern fully reverses, such that participants expect the teacher to select the engineering game for the boy and the story game for the girl, in line with the stereotype, not the students' stated preference. A weaker but still consistent pattern of evidence would be if the pattern attenuates but does not fully reverse, such that participants are less likely to predict that the teacher will select the students' true interests, compared to the Teacher Knowledgeable trials.

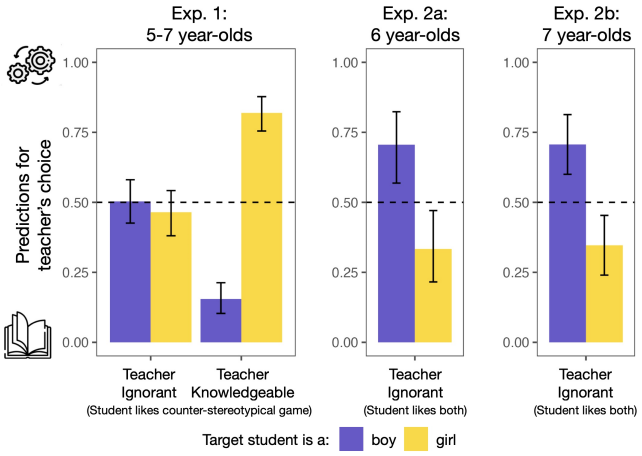

Figure 2: Results from Exp. 1-2 test trials. Each bar depicts proportion of participants predicting that the teacher chose the engineering game in each trial. Error bars show 95% CIs.

Finally, we solicited participants' gender interest stereotypes, so that we could explore how participants' expectations for others' stereotypes are related to their own stereotypes. Participants underwent four stereotype trials about how much boys and girls each like engineering and stories. Based on child-friendly gender stereotype questions from Master et al. (2021), participants were shown a cartoon boy or girl along with the engineering or stories icon and were asked in a two-step format: "Do most girls (boys) like engineering (stories) or do most girls (boys) not like engineering (stories)?" (response options: "like" or "not like"), then "How much do they (not) like it?" ("(not) like a little", "(not) like", or "really (not) like"). Participants responded by clicking thumbs up or thumbs down icons, and responses were converted to a 6-point scale ("Really not like" to "Really like").

Results and Discussion

All analyses were preregistered unless specified otherwise. First, to test our main hypothesis, we ran a mixed-effects logistic regression predicting participants' game predictions, as a function of student gender, teacher knowledge, and their interaction, with random intercepts for participants. As predicted, we found a significant interaction between student gender and teacher knowledge (b=3.27, $p < .001$), showing that participants' predictions of what the teachers chose for boy versus girl students depended on the teacher's knowledge of the student's interests; see Figure 2. We did not find effects of participant age or gender in any of the analyses (when included as fixed effects, with or without an interactive term).

To unpack the interaction between student gender and teacher knowledge on children's choices, we ran mixed-effects logistic regressions within the Teacher Knowledgeable and Teacher Ignorant trials separately, predicting participants' game predictions as a function of student gender with random intercepts for participants. As predicted, in the Teacher Knowledgeable trials, participants were more likely to predict that the teacher chose the engineering game for the girl student than the boy student (b=3.19, $p < .001$), since the teacher knew the students preferred the counter-stereotypical games. In the Teacher Ignorant trials, however, participants did not show a difference in their game predictions for the girl and boy student (b=-.08, $p = .72$).

To better understand why participants did not show a difference in the Teacher Ignorant trials, we ran exploratory analyses examining participant-level responses to the Ignorant trials. The majority of participants either chose the predicted (stereotypical) games for each student (35.2%) or chose the games that reflected the student's actual interests (counter-stereotypical, 33.1%), rather than the same game for both students (both story, 17.9%, both engineering, 13.8%). This pattern of results is different than what would be expected by chance (25% for each pattern, X=10, p=.019, Chi-square), suggesting children *do* have expectations for what the teacher will give to the student, rather than guessing randomly on each trial. We return to this point below.

Next, we explored children's own stereotypes about what girls vs boys are interested in. For each activity (engineering, stories), we ran a mixed-effects linear regression predicting participants' interest ratings as a function of the target gender, with random intercepts for participants. Participants rated boys as more interested in engineering than girls (b=-1.81, $p < .001$) and girls as more interested in stories than boys (b=.86, $p < .001$), demonstrating that participants held gender stereotypes themselves.

Finally, we explored how participants' stereotypes relate to their predictions in the Teacher Ignorant trials. For each participant, we calculated difference scores for how much participants believed each gender was interested in engineering versus stories, generating two difference scores for each participant (i.e., how much they think boys like engineering versus stories, and how much they think girls like engineering versus stories). Then, we ran logistic regressions predicting choice of the engineering game in the Teacher Ignorant trials by these difference scores. The difference scores did not significantly predict participants' choices in the boy trial (b=.01, p=.90) or the girl trial (b=-.08, p=.22).

Here, we found somewhat weak evidence that children expect adults to hold gender stereotypes. Although participants understood that a teacher would select the games that matched a student's interests when the teacher was knowledgeable about the student's interests, they were at chance on whether a teacher ignorant of a student's counter-stereotypical interests would assign them a stereotypical task. Nonetheless, in the Teacher Ignorant trials, most children consistently chose either the predicted (stereotypical) games or the students' actual interests, suggesting children were not guessing randomly. We suspect that children do expect teachers to choose a game for students based on gender stereotypes (when the teacher does not know what the student likes), but have difficulty overriding their *own* knowledge of the stu-

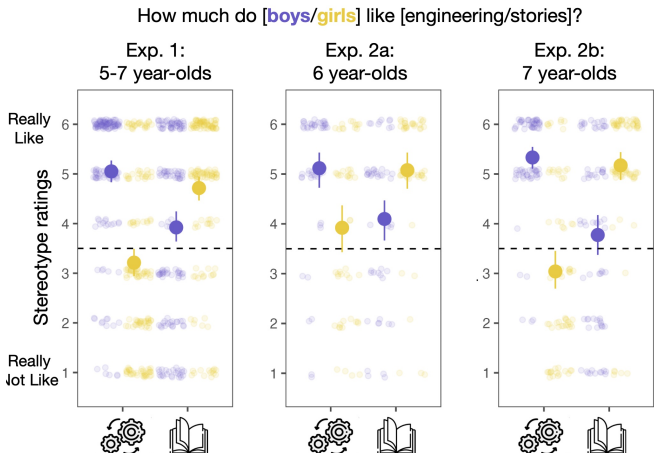

Figure 3: Results from Exp. 1-2 stereotype trials: large points represent mean responses, error bars are 95% CIs, and small points are individual responses.

dent's preference for the counter-stereotypical game.

Experiment 2 tests this possibility with a similar procedure to the Teacher Ignorant trials, except the students now state that they each like *both* engineering and stories. If children expect ignorant adults to hold gender stereotypes, then participants should now predict that the teacher will choose the gender-stereotypical game for each student. However, it is also possible that children simply expect adults to assign different games to boys versus girls, but do not have systematic predictions about which specific games teachers would select for each gender (i.e., they predict one game will be given to one student and the other game to the other student). If this were the case, participants should show the same at-chance predictions as in Experiment 1.

Experiment 2a-2b

Experiment 2 only included the Teacher Ignorant trials and the students stated their interests in both games. Because we were unsure when in development this effect would emerge, we preregistered and ran 6-year-olds (Expermient 2a) and 7-year-olds (Experiment 2b) separately[2].

Methods

Participants Experiment 2a recruited 51[3] 6-year-old participants (Mean age: 6.45 years, SD: .31, Range: 6.08-6.98) and Experiment 2b recruited 75 7-year-old participants (Mean age: 7.47 years, SD: .29, Range: 7.06-7.97) on *Lookit*. Across experiments, parents reported participants' gender as 48.4% girls and 50.8% boys, and .008% (n=1) other, and participants' race as 60.7% White, 9% Asian, 6.6% Asian and White, 5.7% Black, 5.7% White and Hispanic, Latino or Spanish origin, 4.1% Hispanic, Latino or Spanish origin, or 8.2% other categories. Five participants were excluded for participating outside of the US (Exp. 2a: n=1, Exp. 2b: n=4).

Stimuli The stimuli were the same as in Experiment 1, except we switched one of the girl student pictures to a more stereotypically-feminine picture of a girl for clarity.

Procedure The design was identical to Experiment 1, except for the following two changes. First, the procedure only involved the Teacher Ignorant trials. Second, the boy and girl students each stated that they like both engineering and stories. As before, the key test question was which game participants thought the teacher would give to each student (two trials, order counterbalanced). Finally, participants completed the same stereotype questions as in Experiment 1.

Here, we hypothesized that participants would expect teachers to give games to the students based on gender stereotypes about their interests. That is, we predicted that participants would expect that an ignorant teacher would be more likely to give an engineering game to a boy student than to a girl student, even though, as the participant knows, both students actually like both engineering and stories.

Results and Discussion

We ran a mixed-effects logistic regression predicting participants' game predictions, as a function of student gender, and intercepts for subjects. As predicted, we found that both 6- and 7-year-old participants were significantly more likely to predict that the teacher would give the engineering game to the boy student than to the girl student (6-year-olds: b=-1.57, $p < .001$; 7-year-olds: b=-1.51, $p < .001$). Specifically, when the student was a boy, participants selectively predicted that the teacher would choose the engineering game over the story game for him (6-year-olds: 72.34%, p=.003, Binomial Test; 7-year-olds: 70.67%, $p < .001$). When the student was a girl, participants selectively predicted that the teacher would choose the story game over the engineering game for her (6-year-olds: 65.96%, p=.040, Binomial Test; 7-year-olds: 65.33%, p=.011). Thus, even though participants themselves knew that the students would be happy with either game, they predicted that the teacher's game selection would be in line with gender stereotypes. As in Experiment 1, we did not find effects of participant age or gender for the test trials (when included as a fixed effect, with or without an interactive term).

Next, we explored participants' own stereotypes. Similar to Experiment 1, we found that both 6- and 7-year-old participants rated boys as more interested than girls in engineering ($|b|$'s$>=$ .1.4, p's$<$ .001) and girls as more interested than boys in stories ($|b|$'s$>$ .98, p's$<$ .001). Then, for each participant, we computed two difference scores that reflected how much more they believed boys (or girls) liked engineering over stories. As in Experiment 1, we ran logistic regressions predicting participants' choice of the engineering game by these scores, and we did not find, in 2a or 2b, that they predicted participants' choices for the boy ($|b|$'s$>$ .25, p's$>$ .108) or girl trial ($|b|$'s$>$ .11, p's$>$ .327).

[2]Note that we ran the 7 year-olds first, but report the results in chronological age order for ease. Also, we are currently running the experiment with 5 year-olds, but those data are not presented here.

[3]We preregistered 50 6-year-olds and oversampled on accident.

In sum, Experiment 2 showed that 6- to 7-year-old children do indeed expect adults to hold gender interest stereotypes even if they have evidence that these stereotypes are not accurate for novel individuals. Given that the key change between Experiment 1 and 2 is the students' interests, the stronger results in Experiment 2 are consistent with our interpretation that participants may have had difficulty overriding the students' true, counter-stereotypical interests in Experiment 1. Finally, as in Experiment 1, children's own gender stereotypes did not relate to their predictions of the adults' actions.

General Discussion

This work shows that young children expect adults to hold gender stereotypes and represent them separately from their own beliefs. Specifically, we found that 5- to 7-year-olds were more likely to predict that a teacher would give a new student a stereotypical game (e.g., an engineering game to a boy and a story game to a girl) when the teacher did not know (vs. did know) the student's true, counter-stereotypical interest (Exp. 1). When the student liked both games equally, 6- and 7-year-old children predicted that an ignorant teacher would be more likely to give an engineering game to a boy than to a girl (Exp. 2). Taken together, this work shows that by age 6, children expect adults to hold stereotyped beliefs, even when they know that these beliefs are not accurate.

Our work contributes to research on both stereotype reasoning and Theory of Mind. Prior work on children's stereotype awareness has found somewhat mixed results in school-aged children. By using a simple, minimally verbal task (compared to past work, e.g., Kurtz-Costes et al., 2014) and controlling for children's beliefs, we found that, by around the time children enter formal schooling, they already hold expectations that adults have gender stereotypes about novel individuals. Ongoing work is exploring when in development this expectation emerges. Furthermore, work on children's developing Theory of Mind has largely focused on how children *infer* others' beliefs from evidence (e.g., Wellman et al., 2001). Our research poses the novel question of whether children have expectations for others' *prior beliefs*, in the absence of evidence, about new individuals. Indeed, we find that children expect adults to have priors about people's interests in line with gender stereotypes.

Notably, we did not find a correlation between children's expectations of the teachers' stereotypes and their own stereotypes: Participants who provided the predicted responses in Experiments 1-2, that the teacher would give the engineering game to the boy and the story game to the girl, were not more likely to endorse gender interest stereotypes themselves. Given this, one might wonder: Where else might these expectations come from? First, children may observe biases in the adults in their lives, for example, through their explicit statements or more subtle, non-verbal behaviors (e.g., statistical information about what adults tend to praise versus disapprove of, Brey & Pauker, 2019). This may lead children to expect that all adults hold biases, even adults whom they have never met before. In addition, children may pick up on others' gender biases from the institutions that they interact with on a daily basis: Toy stores, TV advertisements, and on (e.g., walk into clothing stores and you will see STEM shirts for boys, but not girls). Given that these biases are seeped into American culture, children might come to expect that most people hold them, even if they themselves do not.

Our work raises the question of whether children expect all adults to hold gender biases. In our experiments, we did not find differences in children's predictions by teacher gender or race, but we may need more trials to detect such effects. One possibility is that children may rely on others' social groups (e.g., their gender or race) to infer which biases they may or may not hold. For example, they might expect people to hold in-group biases and therefore assume that women are more favorable towards women, or more egalitarian (even if this is not actually the case). It is also possible that, with enough evidence, children may come to expect that certain adults *do not* hold gender biases. For example, a child may learn that their parents are egalitarian and not expect them to hold gender biases towards new people but may still expect strangers to hold these biases. More broadly, there are open questions concerning whether children and adults understand how people's social environments (e.g., political affiliation, or neighborhood demographics, Hwang et al., 2021) affect people's stereotypes. Understanding who children expect to hold gender biases is an important area for future research.

Another question for future work is how children's representations of others' stereotypes impacts their own stereotypes and learning behaviors. First, to the extent that children trust the adults in their lives, they may consider their stereotyped beliefs as valid or even prescriptive and take them on as their own. Second, children's assumptions about their teacher's gender interest stereotypes may guide what activities children decide to pursue in the classroom. Given prior work suggesting that both children and adults negatively evaluate children who violate gender stereotypes (Sullivan, Moss-Racusin, Lopez, & Williams, 2018; Skočajić, Radosavljević, Okičić, Janković, & Žeželj, 2020), it is possible that children may want to avoid such judgment or punishment by choosing gender-stereotypical activities. Future work should explore how children's expectations of adults' gender stereotypes impact children's decision-making in the classroom, and importantly, how best to intervene on this process.

Gender stereotypes negatively impact children's everyday decision-making and outcomes in the classroom. Here, we find that 6-to-7-year-old children already expect adults to hold gender stereotypes about students' interests even when children know that these stereotypes are not accurate in this context. Importantly, children's expectations of adults' gender stereotypes may have downstream consequences on their own stereotypical reasoning and behavior in ways that limits their growth. This leaves us with a critical question as scientists and humans: What can we do, as adults, to reverse this expectation for the next generation?

Acknowledgments

We thank Melissa Santos and Natalie Massetti for help with stimuli creation and online recruitment, and all members of the Leonard Learning Lab for helpful feedback and discussions. This work was supported by an NSF SBE Fellowship to MA and a Jacobs Foundation Research Fellowship to JAL.

References

Asaba, M., & Gweon, H. (2022). Young children infer and manage what others think about them. *Proceedings of the National Academy of Sciences*, *119*(32), e2105642119.

Bass, I., Mahaffey, E., & Bonawitz, E. (2021). Do you know what i know? children use informants' beliefs about their abilities to calibrate choices during pedagogy. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 43).

Bian, L., Leslie, S.-J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389–391.

Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. *Advances in child development and behavior*, *34*, 39–89.

Brey, E., & Pauker, K. (2019). Teachers' nonverbal behaviors influence children's stereotypic beliefs. *Journal of Experimental Child Psychology*, *188*, 104671.

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child development*, *82*(3), 766–779.

Cvencek, D., Nasir, N. S., O'Connor, K., Wischnia, S., & Meltzoff, A. N. (2015). The development of math–race stereotypes:"they say chinese people are the best at math". *Journal of Research on Adolescence*, *25*(4), 630–637.

Freeman, N. K. (2007). Preschoolers' perceptions of gender appropriate toys and their parents' beliefs about genderized behaviors: Miscommunication, mixed messages, or hidden truths? *Early childhood education journal*, *34*, 357–366.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child development*, 26–37.

Hwang, H. G., Debnath, R., Meyer, M., Salo, V. C., Fox, N. A., & Woodward, A. (2021). Neighborhood racial demographics predict infants' neural responses to people of different races. *Developmental science*, *24*(4), e13070.

Kurtz-Costes, B., Copping, K. E., Rowley, S. J., & Kinlaw, C. R. (2014). Gender and age differences in awareness and endorsement of gender stereotypes about academic abilities. *European Journal of Psychology of education*, *29*, 603–618.

Master, A., Meltzoff, A. N., & Cheryan, S. (2021a). Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Proceedings of the National Academy of Sciences*, *118*(48), e2100030118.

Master, A., Meltzoff, A. N., & Cheryan, S. (2021b). Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Proceedings of the National Academy of Sciences*, *118*(48), e2100030118.

McKown, C., & Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in middle childhood. *Child development*, *74*(2), 498–515.

Pauker, K., Xu, Y., Williams, A., & Biddle, A. M. (2016). Race essentialism and social contextual differences in children's racial stereotyping. *Child Development*, *87*(5), 1409–1422.

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "michael can't read!" teachers' gender stereotypes and boys' reading self-concept. *Journal of educational psychology*, *107*(1), 186.

Scott, K., & Schulz, L. (2017). Lookit (part 1): A new online platform for developmental research. *Open Mind*, *1*(1), 4–14.

Sigelman, C. K. (2012). Rich man, poor man: Developmental differences in attributions and perceptions. *Journal of Experimental Child Psychology*, *113*(3), 415–429.

Skočajić, M. M., Radosavljević, J. G., Okičić, M. G., Janković, I. O., & Žeželj, I. L. (2020). Boys just don't! gender stereotyping and sanctioning of counter-stereotypical behavior in preschoolers. *Sex Roles*, *82*, 163–172.

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual review of psychology*, *67*, 415–437.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology*, *35*(1), 4–28.

Sullivan, J., Moss-Racusin, C., Lopez, M., & Williams, K. (2018). Backlash against gender stereotype-violating preschool children. *PloS one*, *13*(4), e0195503.

Wellman, H. M., & Banerjee, M. (1991). Mind and emotion: Children's understanding of the emotional consequences of beliefs and desires. *British Journal of Developmental Psychology*, *9*(2), 191–214.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, *72*(3), 655–684. doi: https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child development*, *75*(2), 523–541.