Children's Perceptions of Mechanical Knowledge as a Function of Gender

Lauren Schneider

Cornell University

Author Note

Lauren Schneider, Department of Human Development, Cornell University.

Correspondence may be set to Lauren Schneider, Department of Human Development,

Cornell University, Ithaca, NY 14853; Email: las362@cornell.edu.

Abstract

This study examined children's perceptions of the mechanical knowledge of others as a function of gender. Children ages 3-8 watched videos of male and female informants fixing toys in one of two conditions: either the boy and the girl both succeeded at fixing the toys, or the girl succeeded and the boy failed. The children then answered questions about the informants' abilities to fix other toys. An omnibus ANOVA failed to yield any significant effects of the sex of the informant, though there was a significant effect of condition across age groups and a significant effect of age in the Sally Fixer condition. There is some evidence that nonconformity with gender stereotypes influences children's opinions on informants' play abilities. Further research is needed on the development of children's gender stereotypes, given that children do express gender biases in other contexts.

Keywords: early childhood, mechanical knowledge, gender, epistemic trust

Children's Perceptions of Mechanical Knowledge as a Function of Gender

American mainstream culture is highly gendered. Society's theoretical assumptions about gender, which are not rooted in scientific study, have real-world consequences for both children and adults. Previous research has found that children sometimes use conceptual or cultural biases rather than past reliability to make decisions about who is a reliable informant (Jaswal & Neely, 2006; Harris & Corriveau, 2011). The current research explores whether children believe a relationship exists between the gender of the informant and his or her mechanical knowledge. Specifically, this study looks at whether children will expect that male informants have mechanical knowledge, even when being shown evidence to the contrary. The objectives of this study are: 1) to determine if children track and use knowledge of past mechanical expertise to selectively guide future requests for information, 2) to determine if children rely on gender biases when selecting a preferred informant for new mechanical information, 3) to determine if either of the previous measures changes with age, and 4) to see if past mechanical expertise influences children's predictions about gender stereotyped toy preferences.

Beliefs about gender profoundly affect society's treatment of children, and consequently, their development into adults (Barnett & Rivers, 2004). In infancy, a baby boy may be stuffed into blue sleepers and given trucks and balls to play with, but a baby girl might be dressed in little pink dresses and handled delicately as though she might break (Eliot, 2009). Preschool children's books are full of gender scripts and ideologies, which affect children's role-playing games by prescribing notions of what boys and girls do (Bem, 1981; Bem, 1983). Contemporary coloring books are also replete with gender stereotypes; an analysis of 889 characters in 59 coloring books revealed that males were characterized as more active and were often portrayed as animals, adults, or superheroes, and that both genders were likely to be depicted in a stereotypical way (Fitzpatrick &

McPherson, 2010). Even language is gendered, containing a myriad of assumptions about what boys and girls can or should do (Gelman, 2004). For example, both mothers and their young children are highly accurate at using gendered nouns and pronouns even when a person's gender is presented with the wrong label, and they also use generics to make sweeping statements about gender, e.g. "Boys are good at football" or "That's for boys, not for girls" (Gelman, 2004, p. 60-61).

The divide between male and female is reinforced at every stage of life, with the divides between boys' and girls' play giving way to sexual double standards in adolescence, and to chasms between men and women's career aspirations in adulthood. Language remains gendered; according to Lenton, Sedikides, and Bruder's (2009) latent semantic analysis of a representative sample of college textbooks, "American English reflects and reinforces gender stereotypes regarding gender roles at a level beyond that recognized previously" (p. 269). This means that certain words are more likely to appear in the same context than other words: "Man" is more likely to appear in similar contexts as the words "engineer", "capable" or "independent", while "woman" is more likely to appear alongside words such as "nurse", "emotional", or "shy" (Lenton, Sedikides, & Bruder, 2009, p. 277). In a study of implicit and explicit gender stereotyping of careers, 'engineer' was both implicitly and explicitly related to masculinity (White & White, 2006). Furthermore, research indicates that female students studying math, science, engineering, or technology studies are often overtly or subtly discriminated against in college (Steele, James, & Barnett, 2002), leading them to drop out of their majors. Presently, only 26% of graduate students and 18% of professors in these fields are women (National Science Foundation, 2009).

Given that even adults attribute mechanical abilities to men, young children, who also have strong essentialist beliefs, may also consider mechanical knowledge to be gendered. Currently

there is no research that directly examines this question. There is, however, a significant amount of

work on children's essentialism with regard to gender categories – sometimes referred to as

"gender schemas" – that provide mental prototypes; an ideal version of what a member of a

category is, does, or looks like. Children assume that category membership determines the essence

of what something is (Rhodes & Gelman, 2008), and this assumption is no less true when it comes

to gender. Young children have very strict understandings of gender categories, and therefore have

difficulty understanding prototype exceptions (Eliot, 2009). A good example of the inflexibility of

gender schemas is the relationship between genitals and gender. Because gender and biological sex

are so tightly woven together in the gender binary, children conserve gender across

transformations if they have genital knowledge of the subject (Bem, 1989). For example, they say

a boy dressed as a girl is still a boy because he has a penis; though, as studies with transgender

people have shown, genitals are not the same as gender (Stryker, 2004). Children also believe that

the sex of a baby determines the gender-stereotyped properties the baby will have as he or she gets

older (Taylor, 1996). This research on children's gender essentialism provides an important base

for the current research.

In addition, preschoolers readily infer that individual differences in behavior are tied to

social categories, including gender (Rhodes & Gelman, 2008; Shutts, Banaji, & Spelke, 2007).

Children also make preferential assumptions based on gender, assuming not only that boys like

some things because they are boys, and girls like some things because they are girls, but also that

boys and girls are *supposed* to play with certain toys over other ones (Martin, Ruble, & Szkrybalo,

2002). Furthermore, children assume that category membership, such as gender, involves "role

expectations of prescribed behaviors" (Kalish & Lawson, 2008, p. 588). These findings are

interesting, considering that gender roles are, at least in part, socially constructed (Beal, 1994). In

pre-US Cherokee society, for example, women were the farmers and were included in leadership

(Perdue, 1998), while in present-day India, many men wear skirts and women often wear pants and

a tunic. These examples highlight how roles and behaviors that are considered 'male' in one

culture or time period are not 'male' at all, but can be filled by any person regardless of gender.

Therefore, it is fascinating how readily young children are willing to adopt a society's gendered

notions regarding individuals' behavior.   The social construction of stereotype development in

children is the focus of developmental intergroup theory (Bigler & Liben, 2007), which suggests

that stereotypes develop early in childhood and are largely under environmental control.

     The current research hopes to shed light on the development of gender stereotypes about

mechanical knowledge. The basis of this study lies in integrating the work on gender stereotype

development with children's learning from reliable informants. The research on children's learning

from informants has exploded in the last ten years. In general, children are very good at figuring

out which people are good sources of information. In toddlerhood, children become attuned to

nonverbal cues from adults about who accurate informants are, and they therefore prefer to learn

from those informants who seem knowledgeable (Birch, Akmal, & Frampton, 2010). By age four,

children mistrust ignorant and inaccurate informants in both verbal and nonverbal domains

(Koenig & Harris, 2005), and they also track and use past accuracy to guide their learning (Birch,

Vauthier & Bloom, 2008). In addition, children differentiate between causal expertise and other

types of expertise when presented with informants who know some things but not others (Kushnir

& Vredenburgh, 2010, in preparation).

     Not only are children able to discriminate reliable informants based on knowledge, they are

also capable of rejecting the common stereotype equating adulthood with greater expertise. For

example, if an adult proves to be ignorant, children will learn word labels from an expert child

rather than from the ignorant adult (Jaswal & Neely, 2006). In a review of the literature on this topic, Harris and Corriveau (2011) argue for the existence of two heuristics that influence children's epistemic trust: 1) Children trust based on an informant's past accuracy, and 2) Children trust based on an informant's cultural standing. This has great bearing on the current research. In our culture, adults are (usually) deemed to be more trust worthy than children (Jaswal & Neely, 2006), and as described above, men are deemed more trustworthy than women in mechanical domains. Because this stereotype and others about gender are so powerful, the question arose: Will children's beliefs about gender prove to be more powerful than what they observe about the reliability of informants? Will they consider men to be more reliable than women, particularly in a stereotypically masculine field as mechanical knowledge? The current research hopes to answer these questions.

     The overall goal of this study is to probe whether or not children associate mechanical knowledge with being male, and whether they use that bias to decide who is a trustworthy informant. I predict that by age four, children have a gender bias towards males in the realm of mechanical knowledge. I also predict that by age seven, children's gender bias will be more pronounced. As I have described above, gender stereotypes permeate every life stage. Therefore, the longer children have been alive, the more likely they are to have internalized stereotypes about the mechanical abilities of boys and girls. In addition, children's ability to track and use information about an informant's accuracy improves with age (Fitneva & Dunfield, 2010; Einav & Robinson, 2010), so it would take an override of ability by stereotype for a child to choose the male informant. Ideally, children who see videos of men and women successfully fixing toys should pick one at chance, but I hypothesize that they will pick the males over the females more often. This is because I expect gender biases in the absence of information to the contrary, as

Jaswal and Neely (2006) found ageist biases at baseline. Conversely, I predict that children who

see videos of successful females and unsuccessful males will pick the females, but at lower rates

than the baseline group picked males. This study will hopefully shed light on the development of

children's gender stereotypes and provide a basis for future work on this topic.

Method

Participants

Participation was limited to preschool and elementary school children between the ages of

three and eight. Participants were recruited from preschools in Ithaca, NY, and New York, NY,

and elementary schools in Ithaca, NY and Syracuse, NY. Only children whose parents permitted

them to participate in the study were played the research game. Altogether, 48 children

participated in the study: 32 3-5 year olds (Range = 3.66-5.35; M = 4.36; SD = 0.36; 18 males) and

16 6-8 year olds (Range = 6.03-8.57; M = 7.11; SD = 0.89; 9 males). Within the preschool group,

25 of the participants were white, 3 were Hispanic/Latino, 1 was Black, 2 were Asian/Pacific

Islander, and 1 was multi-racial. Within the elementary school group, 12 of the participants were

white, 1 was Black, and 3 were Asian/Pacific Islander.

Materials

The experimental materials included four novel "broken" toys, two unfamiliar objects, and

videos of young adult confederates (one male and one female) attempting to fix toys. The

confederates were Cornell students never seen before by any of the children. Both informants had

a stereotypically gendered appearance: The boy had short hair, slight stubble on his face, chest

hair, a deep voice, masculine features, and was wearing a blue shirt; the girl had long blonde hair,

breasts, a high voice, feminine features, and was wearing a red shirt. There was only one informant

of each gender because Jaswal and Neely (2006) found that one informant of each age was significant to find an ageist effect. We anticipated, therefore, that it would require only one informant of each gender to find a sexist effect. Each child saw each informant multiple times throughout the course of the experiment.

Within the videos, the confederates attempted to fix the following toys during the history phase: A toy boat, a toy turtle, a set of toy car keys, and a toy microphone. All the toys made noise when they were "fixed". Each informant attempted to fix two of the four toys. The four novel "broken" toys that were shown to the children during the test phase were a toy car, a toy cell phone with wheels, a toy camera, and a toy that sang the alphabet. The two unfamiliar objects were a refillable sponge and an egg slicer. In the test phase, the children saw three more videos of each informant offering explanations about why two out of four novel toys were broken, as well as an explanation of what the egg slicer was for.

Apparatus

Videos were played using QuickTime on a black Mac Book running OS 10.6.5. The computer was placed on a table in front of the child's face. The camera used to record each session was a Sony Handycam. The camera was placed in such a way as to record the child's face and what he or she was saying.

Procedure

The researcher visited the preschool and played the game with the child in a quiet, open space visible to the staff of the testing site. The child would watch the videos and the researcher would ask the questions. The answers to the questions make up the data set.

The experimental design for this study was based on Jaswal and Neely's (2006) study design, where they paired videos of a reliable and/or unreliable adult with videos of a reliable

and/or unreliable child. Similarly, in this study, participants were randomly assigned to view one

of the following two pairs of videos in the history phase, which corresponded to the two

conditions: A boy who could fix toys and a girl who could fix toys (baseline/Both Fixers

condition), or a girl who could fix toys and a boy who could not (anti-stereotype/Sally Fixer

condition). After each video segment was presented, the experimenter would say, "S/he fixed it!"

when the informant was successful, or "He didn't fix it!" when the boy failed to fix the toy. After

the history phase, before moving on to the test phase, children were asked to recall if the

informants had successfully fixed the toys. The purpose of the memory question was to check if

the children paid attention to the videos. Both boy/girl order and toy type were counterbalanced in

a Latin square.

In the test phase of each condition, children were presented with four novel "broken" toys

(a car, a cell phone with wheels, an alphabet toy, and a camera) and two novel artifacts (a refillable

sponge and an egg slicer). Each object had a question associated with it. The order in which toys

and the test objects were presented was all counterbalanced in a Latin square design. For two of

the toys, the children were asked, "Who should I ask to help me fix this toy? Should I ask Bill, or

should I ask Sally?" For two of the toys, the children watched as Sally and Bill each offered an

explanation for why the toy was broken, and then asked to endorse one of the explanations:

"What's wrong with this toy? Is it the motor, or the batteries?" and "What's wrong with this toy?

Is it the wires, or the gears?" The purpose of having both "ask" and "endorse" questions is that

Koenig and Harris (2005) found that four-year-olds consistently answer both of these types

questions based on the past accuracy of the informant.

The purpose of the control questions about the refillable sponge and the egg slicer was to

control for informant preference. Kushnir et. al. (2010) found that children do not generalize an

informant's knowledge of causal mechanisms to knowledge of labels, and vice versa. Based on this research, children should answer these questions at chance. For the refillable sponge, children were asked who would know the name of the toy: "I don't know what this thing is called. Who should I ask? Should I ask Bill, or should I ask Sally?" For the egg slicer, the children watched as Sally and Bill each offered an explanation of what the object was for, and then asked to choose one of the explanations: "What is this for? Is it for making music, or is it for clapping?" Boy/girl order was the same as it had been in the information phase. The script for the procedure can be viewed in the Appendix.

Following the completion of the test phase, a subset of preschool children were also asked about both informants: 1) Is Bill/Sally good at playing with trains or playing with dolls?, 2) Is Bill/Sally good at playing baseball or dancing ballet? and 3) Is Bill/Sally good at playing with blocks or playing in the kitchen? A diagram of the toys accompanied the questions. Four children in the baseline condition were asked these questions, and six children in the anti-stereotype condition were asked these questions. The script for this part of procedure can be viewed in the Appendix.

Following the completion of the test phase, a subset of elementary school children were also asked about both informants: 1) Do you think that Bill/Sally could fix a car engine?, 2) Do you think that Bill/Sally could fix a computer? and 3) Do you think that Bill/Sally could fix a bicycle? A diagram of the objects accompanied the questions. Three children in the baseline condition were asked these questions, and seven children in the anti-stereotype condition were asked these questions. The script for this part of procedure can be viewed in the Appendix.

Coding

For each question in the test phase, the children's answers were scored as 0 for Bill and 1 for Sally. "Both", "maybe", and "I don't know" answers were coded as 0.5 (chance), but these answers were very uncommon. Following preliminary analysis each child received a score out of 4, where 0 was equivalent to choosing Bill 4 times, and 4 was equivalent to choosing Sally 4 times.

Results

Sixteen preschool children saw Bill presented before Sally in the history and test phases of the videos, and sixteen preschool children saw Sally presented before Bill in the history and test phases of the videos. Due to a lack of participation from elementary schools, the sixteen elementary school children only saw Bill presented before Sally in the history and test phases of the videos. The first analysis, therefore, tested for informant order effects between questions. A McNemar's test comparing responses to the first "ask" question and the second "ask" question found no significant difference in responses to the two questions ($p = 1$). Two children who answered at chance were not included in this analysis. A second McNemar's test was conducted comparing answers to the first "endorse" question to the second "endorse" question and found no significant difference in responses to the two questions ($p = 1$). A third McNemar's test was done to assess any significant differences between answers to the "motor/batteries" endorse question and the "wires/gears" endorse questions and found no significant difference between answers to the two questions ($p = 0.189$).

An omnibus ANOVA comparing age group (preschool vs. elementary school) to condition (Both Fixers vs. Sally Fixer) to a within subjects factor of question type ("ask" vs. "endorse") was conducted to check for differences in question type and for any main effects of age or condition.

The ANOVA revealed no within-subjects difference in question type ($F(1, 48) = 1.464$, $p = 0.333$), and no interaction effects between the within-subjects factor "question type" and age ($F(1, 48) = 0.969$, p = 0.268) or condition ($F(1, 48) = 0.163$, $p = 0.689$). There were significant effects of condition ($F(1, 48) = 20.023$, $p < 0.001$) and age group ($F(1, 48) = 4.977$, $p = 0.031$), but no interaction effects between age and condition ($F(1, 48) = 0.553$, $p = 0.461$). A graphic depiction of these results can be seen in Figure 1.

To analyze the age effect, two *t*-tests were conducted: One comparing the preschoolers in the baseline condition to the elementary school children in the baseline condition, and one comparing the preschoolers in the Sally Fixer condition to the elementary school children in the Sally Fixer condition. The *t*-test comparing the two age groups in the baseline condition revealed no significant differences ($t(22) = -0.888$, $p = 0.384$). However, the *t*-test comparing the two age groups in the Sally Fixer condition revealed a significant difference between preschooler's and elementary school children's responses ($t(22) = -2.724$, $p = 0.012$). Explanations for this result can be found in the Discussion section.

A final analysis was done to check for effects of school location. A univariate ANOVA with school location as a random factor revealed no significant differences in answers to any of the questions between the different school locations ($F(3, 48) = 1.603$, $p = 0.202$), and no significant interaction effect for condition and school location on answers ($F(3, 48) = 0.825$, $p = 0.488$).

As stated in the Procedure section, all the children were asked the same control questions: "Who should I ask what this is called?" about the name of a refillable sponge, and "What is this for? Is it for making music or is it for clapping?" about the function of an egg slicer. Though one four-year-old child did answer, "It's for eggs," the children appeared to choose their informants for these questions at chance. Preliminary descriptive statistics show that in the baseline condition,

about half of the preschoolers chose the same informant for each question (nine out of sixteen) and

the other half chose different informants for each question. In the Sally Fixer condition, exactly

half of preschoolers chose the same informant for each question (eight of out sixteen) and the other

half chose different informants. For elementary school children in both the baseline and the anti-

stereotype/Sally Fixer condition, seven out of eight participants gave the same answer for both

types of questions. An ANOVA comparing answers to label and function question by condition

and age group revealed that there were no condition or age effects for answers to the control

questions (Condition/Label: $F(1, 48) = 0.169$, $p = 0.683$; Condition/Function: $F(1, 48) = 1.108$, $p =$

0.298; Age Group/Label: $F(1, 48) = 2.697$, $p = 0.108$; Age Group/Function: $F(1, 48) = 1.595$, $p =$

0.213. Preschool children chose a labeler at chance, and were a bit more likely (though not

significantly) to choose Sally as their function-explainer (Figure 2), while elementary school

children were slightly more likely to choose Bill as both their labeler and their function-explainer

(Figure 3). Overall, both groups of children choose answers to these control questions at chance.

Indicators of a gender bias were found as well. As stated in the Procedure section, some

preschool children were asked questions about Sally and Bill's ability to play with other kinds of

toys (Appendix 2). Each informant was assigned a score out of 3 for boy toy attributions and a

score out of 3 for girl toy attributions. A between-subjects ANOVA was conducted to check for

differences in answers to questions about Bill and Sally's playing abilities between conditions and

found no significant differences between conditions on any of the attributions (Bill Boy Toys: $F(1,$

$10) = 2.560$, $p = 0.148$; Bill Girl Toys: $F(1, 10) = 2.560$, $p = 0.148$; Sally Boy Toys: $F(1, 10) =$

0.348, $p = 0.572$; Sally Girl Toys: $F(1, 10) = 0.721$, $p = 0.420$).

Though the results of the ANOVA were not statistically significant, they are indicative of a

possible trend towards gender bias. Means, standard deviations, and 95% confidence intervals

were also reported for each gendered toy attribution score for each informant. A graphic depiction

of the scores for these gendered toy attributions can be viewed in Figure 4. When Bill and Sally are

both fixers, Sally was assigned both boy toys ($M = 1.25$, $SD = 0.547$, 95% CI [-0.012, 2.512]) and

girl toys ($M = 1.5$, $SD = 0.608$, 95% CI [0.098 2.902]), but Bill was only assigned only boy toys

($M = 3$, $SD = 0.323$, 95% CI [2.256, 3.744]), and no girl toys ($M = 0$, $SD = 0.323$ 95% CI [-.744,

.744]). However, when only Sally is the fixer, the children still attribute both boy and girl toys to

Sally (Sally's Boy Toy Attributions: $M = 0.833$, $SD = 0.447$, 95% CI [-0.197, 1.864]; Sally's Girl

Toy Attributions: $M = 2.167$, $SD = 0.497$, 95% CI [1.022, 3.312]), but a small number of children

attribute girl toys to Bill as well; his mean girl toys score goes up to 0.667 ($SD = .264$ 95% CI

[.059, 1.274]), and his boy toys score goes down ($M = 2.333$, $SD = 0.264$, 95% CI [1.726 2.941]).

Because the children consistently attributed both boy's and girl's toys to Sally, each

informant's toy preferences were examined without taking condition into account. A t-test

comparing each informant's stereotyped attributions mean (Sally Girl Toys and Bill Boy Toys) to

the expected score of 3 (all stereotyped answers) reveals that Sally is more likely overall to be

given non-stereotypical attributions than Bill is (Sally: $t(9) = 2.905$, p = 0.017); Bill: $t(9) = -1.809$,

p = 0.104). Given the small number of participants in each condition, χ^2 tests are inappropriate

here. Overall, the children were more likely to make gendered attributions to the informants than

non-gendered ones, with a tendency to stereotype Sally less strongly than Bill, but the results when

condition was taken into account were not statistically significant.

Some elementary school children were asked questions about Sally and Bill's ability to fix

other mechanical objects. Each informant was assigned a score out of 3 based on children's

responses to questions about their abilities to fix three real-world objects. A between-subjects

ANOVA was conducted to check for differences in answers between conditions. Because of the

small number of participants in the baseline condition for this question, the results were not statistically significant (Bill Ability: $F(1, 10) = 2.778$, $p = 0.134$; Sally Ability: $F(1, 10) = 0.646$, $p = 0.445$), but it is plausible that there is a trend developing. Means, standard deviations, and 95% confidence intervals of each kind of score reveal that when both informants can fix the toys, children were equally likely to select them to fix real-world objects (Bill: M = 2, SD = 0.499, 95% CI [0.925, 3.075]; Sally: M = 2, SD = 0.520, 95% CI [0.800, 3.200]). However, in the Sally Fixer condition, Sally is rated more highly than Bill (though lower than she was in the baseline condition). Bill's mean score in the Sally Fixer condition was 1.071 ($SD = 0.305$, 95% CI [0.338, 1.775]), while Sally's mean score was 1.5 ($SD = 0.341$, 95% CI [0.714, 2.286]). It is possible that this trend would become more salient if more children were tested. Given the small number of participants in the baseline condition for this question, χ^2 tests are inappropriate here as well. A graphic depiction of the mean scores for these real-world fixing ability questions can be viewed in Figure 5.

Discussion

At the start of this study, I predicted that by age four, children will demonstrate gender bias towards males in the realm of mechanical knowledge, and that they will pick males over females more often. I also predicted that children who see videos of successful females and unsuccessful males will pick the females, but at lower rates than the baseline group picked males. Overall, I expected a gender bias in the baseline condition for both groups, with elementary school children's bias being more pronounced. However, the children demonstrated that, at least in this scenario, that they track the given information only, which is consistent with past research (Birch et. al.,2008; Koenig & Harris, 2005). There was a significant main effect of condition across age

groups, indicating that at least for this protocol, preschool and elementary school children select a preferred fixer based on past information given. There was also a significant effect of age on answers in the Sally Fixer condition. Answers in the baseline condition in both age groups were at chance, but elementary school children were much more consistent than preschoolers at choosing Sally as their preferred informant in the Sally Fixer condition. This is likely an effect of elementary school children's improved ability at tracking an informant's past accuracy and sensitivity to an informant's mistakes (Einav & Robinson, 2010) .

Despite not having gender effects for the target questions, children still expressed gender biases when asked about other possible scenarios. In fact, the children often spontaneously made gendered judgments during the game, saying things like, "My dad can fix it because he is a boy," or, "Ballet is only for girls." Children in both age groups and of both genders made statements like these. As stated in the results section, when Bill and Sally are both fixers, preschool children attribute both boy and girl toys to Sally, but only boy toys to Bill. When only Sally is the fixer, the children still attribute both boy and girl toys to Sally, but a small number of children attribute girl toys to Bill as well. This is interesting, considering what Rhodes et. al. (2008) reported about preschoolers' inferences about play abilities: Children "draw conclusions about stable gender-linked differences on the basis of single instances of relative failure at playing with a toy, which suggests possible processes that may contribute to the early development of children's academic gender stereotypes and achievement behaviors" (p. 971). In both conditions, Sally's femininity didn't stop the children from consistently attributing both boy and girl toys to her, but only Bill's failure to fix toys led to girl toys being attributed to him. It is possible that Bill's inability to fix toys has rendered him less manly, or more feminine, in the eyes of the children. This may be related to the notion that when a man fails to express masculinity in a socially sanctioned way, he

is labeled with homophobic slurs, which are also sexist slurs, since many homophobic slurs imply

that the man in question "lack[s]…masculinity" or "act[s] like a girl" (Plummer, 2001, p.18).

These results hint at the possibility that nonconformity with masculine tropes is much more

notable to children than nonconformity with feminine tropes, which is pertinent to a richer

understanding of Developmental Intergroup Theory (Bigler & Liben, 2007). Because children are

learning at an early age such a narrow definition of masculinity (Bem, 1984; Gelman, 2004;

Fitzpatrick & McPherson, 2010), though they know almost nothing about Bill as a person, it is

possible that their understanding of masculinity can be changed using environmental messages

about how *people* have different abilities, not men and women

Though Sally's girl toy attributions were also slightly higher in the Sally Fixer condition

than they were in the baseline condition, she was assigned boy and girl toys in both conditions.

This may be because it is considered acceptable nowadays for a girl to do "boy" activities. It is

also possible that seeing a woman handling tools and fixing toys was unusual enough to the

children that they became less likely to attribute only girl's toys to someone who had just broken

down a stereotype about women. It may also be that the increase in girl toy attributions to Sally is

due to the possibility that the difference in ability between Sally and Bill made the children also

notice their gender difference. This is a bias that may actually be mitigated by time in school.

When some elementary school children were asked questions about Sally and Bill's ability to fix

other mechanical objects, there was no difference between Sally's and Bill's perceived abilities in

the baseline condition. However, only three children in the baseline condition were asked about the

informants' real-world mechanical object fixing abilities, making these results not generalizable.

When Sally is the only fixer, her ability to fix other objects is rated more highly than Bill's ability.

Because seven children in the anti-stereotype condition were asked about real-world mechanical

object fixing ability, these results are a bit more meaningful. This is further evidence of children's

epistemic tracking: When Sally and/or Bill demonstrate the ability to fix toys, the children may

think the informant(s) can fix other objects as well. More elementary school children need to be

tested to determine whether or not this trend is statistically significant.

An interesting finding was the absence of condition effects on the control questions. In the

group of preschoolers, no matter what the condition, children chose an informant for the labeler at

chance, but were slightly more likely to choose Sally as the function-explainer. However, for the

elementary school children, children were likely to choose an informant at chance, no matter the

question or condition. Even when Sally is the obvious fixer, children in both age groups do not

generalize about her other abilities. This is consistent with past research on children's domain-

specific epistemic trust (Kushnir et. al., 2010): Just because an informant is knowledgeable in one

domain doesn't mean a child assumes that he or she is knowledgeable in another domain.

This study had several strengths. First, the methodology was based on past successful

experimental designs (Koenig & Harris, 2005; Jaswal & Neely, 2006). Second, it was able to show

a significant main effect of condition: Children in both used the past mechanical accuracy of both

informants to guide their future choices. Third, it showed an effect of age related to an

improvement in tracking ability, which is consistent with past research on age differences in

selective trust (Fitneva & Dunfield, 2010; Einav & Robinson, 2010). Fourth, it confirmed past

findings about children's domain-specific information tracking: Children do not generalize

mechanical ability or lack thereof to knowledge about other domains. Overall, the main findings of

this study are consistent with past research on children's tracking of informant's accuracy.

Though it had some strengths, this study was certainly not without shortcomings. First,

there were half as many elementary school children as there were preschoolers, due to a lack of

participation from elementary schools, and consequently, the elementary school children only saw

videos where Bill was presented first. Showing half the children Sally first would have meant that

there would have been only four children in each condition and informant presentation order

combination. Further studies on elementary school children will need to be preceded by more

active and tenacious recruitment at all local schools. A second weakness was the lack of

racial/ethnic diversity in the sample, and the lack of information about children's socioeconomic

status. Further research would need to specifically collect this information in order to check for

racial and socioeconomic cohort effects.

   It is possible, though highly unlikely, that within the six-month data collection period,

some of the children may have seen "Sally" previously at their test site, and as children are more

likely to trust a familiar informant (Corriveau & Harris, 2009), they may have been more prone to

select Sally. The student who played Sally confirmed that this possibility was unlikely (K. Braun,

personal communication, May 11, 2011). Any repetition of this study should use non-researcher

volunteers, instead of volunteers who may at any point be conducting different research in the

same places.

   Another possibility is that the four-year-olds required more examples of Bill and Sally

fixing or failing to fix toys in the history phase, given that four-year-olds may require more

accumulation of information before being able to make a judgment not at chance about an

informant (Fitneva & Dunfield, 2010). Children may also need to be exposed to multiple

informants of each gender, rather than just one, in order to make a sexist judgment about the

abilities of those informants. I also used the same protocol for both age groups. It is possible that I

should have used a different protocol with the older group, since the one devised may have been

too easy for them. It is also possible that I needed to be more explicit about the gender of the

confederates in the video, or prime the children to pay close attention to the gender of the confederates.

The goal for any further studies would be to find out when gender stereotypes about mechanical ability emerge in development. Directions for further research may include in-depth interviews with children about what they think about gender, gender roles, and ability, as well as repeating the study with realistic objects instead of toys. Another direction for research would be to see if children in non-traditional families (e.g. gay parent families, single-parent families) have fewer or different gender biases than children in traditional two-parent heterosexual families. A third area for further research would be to experiment with giving children information about confederate's abilities in domains other than the target domain, to see if children continue to generalize e.g. making Bill a trustworthy labeler and making Sally a non-trustworthy labeler.

This study is a building block towards understanding the development of gender bias in children. By understanding the causal mechanisms of the development of prejudice, policymakers, educators, and others (including parents) can develop strategies to mitigate the development of prejudicial biases and the harmful effects of sexism. It is important that children receive egalitarian messages in their youth, so that in their adult lives they might participate in a society where they are not judged and do not judge others based on gender stereotypes.

References

Barnett, R., & Rivers, C. (2004). *Same Difference*. New York, NY: Basic Books.

Beal, C. R. (1994). *Boys and girls: The development of gender roles.* New York, NY: McGraw-Hill.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88,* 354-64.

Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs: Journal of Women in Culture and Society, 8*, 598-616.

Bem, S. L. (1989). Genital knowledge and gender constancy in preschool children. *Child Development, 60,* 649-662.

Bigler, R. S., Arthur, A. E., Hughes, J. M., & Patterson, M. M. (2008). The politics of race and gender: Children's perceptions of discrimination and the U. S. presidency. *Analyses of Social Issues and Public Policy, 8*, 83-112. doi:10.1111/j.1530-2415.2008.00161.x

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science, 16*, 162-166.

Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year olds spontaneously use others' past performance to guide their learning. *Cognition*, *107*, 1018-1034. doi:10.1016/j.cognition.2007.12.008

Birch, S. A. J., Akmal, N., & Frampton, K. L. (2010). Two-year-olds are vigilant of others' non-verbal cues to credibility. *Developmental Science, 13*(2), 363-369.

doi:10.1111/j.1467-7687.2009.00906.x

Corriveau, K., & Harris, P. L. (2009). Choosing your informant: weighing familiarity and recent

accuracy, Developmental Science,12(3), 426-437.

doi:10.1111/j.1467-7687.2008.00792.x

Einav, S., & Robinson, E. J. (2010). Children's sensitivity to error magnitude when evaluating

informants. *Cognitive Development, 25*(3), 218-232.

doi:10.1016/j.cogdev.2010.04.002

Eliot, L. (2009). *Pink brain, blue brain*. New York, NY: Houghton Mifflin.

Fitneva, S. A., & Dunfield, K. A. (2010). Selective information seeking after a single encounter.

Developmental Psychology, 46(5), 1380-1384.

doi:10.1037/a0019818

Fitzpatrick, M. J., & McPherson, B. J. (2010). Coloring within the lines: Gender stereotypes in

contemporary coloring books. *Sex Roles, 62,* 127-137.

doi:10.1007/s11199-009-9703-8

Gelman, S. A., Taylor, M G., & Nguyen, S. (2004). Mother-child conversations about gender:

Understanding the acquisition of essentialist beliefs. *Monographs of the Society for

Research in Child Development, 69(1, Serial No. 275)*.

Harris, P. L., & Corriveau, K. H. (2011). Young children's selective trust in informants.

Philosophical Transactions of the Royal Society of Biological Sciences, 366, 1179–1187.

doi:10.1098/rstb.2010.032

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past

reliability over age when learning new words. *Psychological Science*, *17*, 757-758.

Kalish, C. W., & Lawson, C. A. (2008). Development of social category representations: Early

appreciation of roles and deontic relations. *Child Development, 79*(3), 577-593.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers.

Child Development, *76*, 1261-1277.

Kushnir, T. K., & Vredenburgh, C., 2010, in preparation

Lenten, A. P., Sedikides, C., & Bruder, M. (2009). A latent semantic analysis of gender-stereotype

constancy and narrowness in American English. *Sex Roles, 60,* 269-278.

doi:10.1007/s11199-008-9534-z

Liben, L. S., & Bigler, R. S. (2008). Developmental gender differentiation: Pathways in

conforming and nonconforming outcomes. *Journal of Gay and Lesbian Mental Health, 12*,

95-119.

doi:l0.13(K)/J529vl2n01 07

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender

development. *Psychological Bulletin, 128,* 903 – 933.

National Science Foundation. (2009). *Women, minorities, and persons with disabilities in science

and engineering: 2007.* Arlington, VA: National Science Foundation.

Perdue, T. (1998). *Cherokee women: Gender and culture change, 1700-1850.* Lincoln, NE:

University of Nebraska Press.

Plummer, D. C. (2001). The quest for modern manhood: masculine stereotypes, peer culture and

the social significance of homophobia. *Journal of Adolescence*, 24, 15–23.

doi:10.1006/jado.2000.0370

Rhodes, M., & Brickman, D. (2008). Preschoolers' responses to social comparisons involving

relative failure. *Psychological Science, 19*, 968-972.

Rhodes, M., & Gelman, S. (2008). Categories influence predictions about individual consistency. *Child Development, 79*, 1270 – 1287.

Shutts, K., Banaji, M.R., & Spelke, E.S. (2007, October). Social categories guide young children's preferences for novel objects. Paper presented at the biennial meeting of the Cognitive Development Society, Santa Fe, NM.

Steele, J., James, J. B., & Barnett, R. C. (2002). Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. *Psychology of Women Quarterly, 26,* 46 –50.
doi:10.1111/1471-6402.00042

Stryker, S. (2004). Transgender. (2004). glbtq: an encyclopedia of gay, lesbian, bisexual, transgender, and queer culture. Chicago, IL: glbtq, Inc.

Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development, 67*, 1555-1571.

Weisgram, E. S., Bigler, R. S., & Liben, L. S. (2010). Gender, values, and occupational interests among children, adolescents, and adults. *Child Development, 81*, 778–796.

White, M. J., & White, G. B. (2006). Implicit and explicit gender stereotypes. *Sex Roles, 55,* 259-266.
doi:10.1007/s11199-006-9078-z

Appendix

1. Protocol

Version A: Bill and Sally are both Fixers (baseline/Both Fixers condition)

[On screen are stills of a boy and a girl. The videos the children watch are counterbalanced by sex and by toy]

Researcher: These are my friends, Bill and Sally. I gave Bill (or Sally) a broken toy to fix.

[On screen, child watches a video of Bill or Sally. Bill or Sally fixes the toy].

Researcher: She/he fixed it!

Researcher: OK! I gave Bill (or Sally) another broken toy to fix.

[On screen, child watches a video of Bill or Sally. Bill or Sally fixes the toy].

Researcher: She/he fixed it!

Researcher: OK! [shows stills again] I also gave Sally (or Bill) a broken toy to fix.

[On screen, child watches a video of Sally or Bill. Sally or Bill fixes the toy].

Researcher: She/he fixed it!

Researcher: OK! I gave Sally (or Bill) another broken toy to fix. [On screen, child watches a video of Sally or Bill. Sally or Bill fixes the toy].

Researcher: She/he fixed it!

Version B: Sally is the Fixer and Bill is not (anti-stereotype/Sally Fixer condition)

[On screen are stills of a boy and a girl. The videos the children watch are counterbalanced by sex and by toy, as per counterbalancing sheet]

Researcher: These are my friends, Bill and Sally. Say hi to Bill and Sally. [pause] I gave Bill (or Sally) a broken toy to fix.

[On screen, child watches a video of Bill or Sally. If it's Bill, he does not fix it. If it is Sally, she does].

Researcher: She fixed it/He didn't fix it!

Researcher: OK! I gave Bill (or Sally) another broken toy to fix.

[On screen, child watches a video of Bill or Sally. If it's Bill, he does not fix it. If it is Sally, she does]
.
Researcher: She fixed it/He didn't fix it!

Researcher: OK! [shows stills again] I also gave Sally (or Bill) a broken toy to fix.

[On screen, child watches a video of Sally or Bill. If it's Bill, he does not fix it. If it is Sally, she does].

Researcher: She fixed it/He didn't fix it!

Researcher: OK! I gave Sally (or Bill) another broken toy to fix.

[On screen, child watches a video of Sally or Bill. If it's Bill, he does not fix it. If it is Sally, she does].

Researcher: She fixed it/He didn't fix it!

Test Phase:

Memory Questions:

Researcher: Now I have a question. Did Bill fix the toys?

Child: Yes/No

Researcher: Did Sally fix the toys?

Child: Yes/No.

Researcher: OK! I have another broken toy here. [produces broken toy, shows stills on the screen]. Who should I ask to help me fix the toy? Should I ask Sally or Bill? (Should I ask Bill or Sally?)

Child: [Name]

Researcher: Why should I ask [Name] to help me fix the toy?

Child: [Answers]

Researcher: OK! I have another toy here. [produces broken toy, shows stills on the screen]. This toy is not working. Let's see what Bill and Sally think about this toy.

[Onscreen, 2 videos, counterbalanced. One, Bill or Sally examines the toy and says: "I think the motor has stopped moving. Two, Bill or Sally examines the toy and says, "I think this toy is out of batteries."]

Researcher: What is wrong with this toy? Is it the motor or the batteries?

Child: Motor/batteries.

Researcher: OK! I have this thing. I don't know what it is called. Who should I ask what this is called/for? Should I ask Bill, or should I ask Sally?

Child: [Name]

Researcher: OK! I have another broken toy here. [produces broken toy, shows stills on the screen]. Who should I ask to help me fix the toy?

Child: [Name]

Researcher: Why should I ask [Name] to help me fix the toy?

Child: [Answers]

Researcher: OK! I have another toy here. [produces broken toy, shows stills on the screen]. This toy is not working. Let's see what Bill and Sally think about this toy.

[Onscreen, 2 videos, counterbalanced. One, Bill or Sally examines the toy and says: "I think the wires are disconnected. Two, Bill or Sally examines the toy and says, "I think the gears are out of sync."]

Researcher: What is wrong with this toy? Is it the wires or the gears?

Child: Wires/Gears.

Researcher: OK! I have this thing. I don't know what it is for. Let's see what Bill and Sally think about this thing:

[Onscreen, children see videos of Bill and Sally, counterbalanced. They each examine the object. One says, "I think this thing is for making music" and plays with the metal bits. The other says, "I think this thing is for clapping".]

Researcher: What is this for? Is it for making music, or is it for clapping?

Child: [Answers]

Researcher: OK! Thank you so much for playing with me, [child's name]. You did a great job.

2. Preschooler's Toy Questions

Children were shown the diagram below and asked:

Is Bill/Sally good at playing with trains or playing with dolls?
Is Bill/Sally good at playing baseball or dancing ballet?
Is Bill/Sally good at playing with blocks or playing in the kitchen?
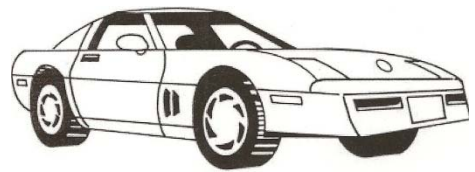
3. Elementary School Children's Object Questions

Children were shown the diagram below and asked:

Do you think that Bill/Sally could fix a car engine?
Do you think that Bill/Sally could fix a computer?
Do you think that Bill/Sally could fix a bicycle?

Figures
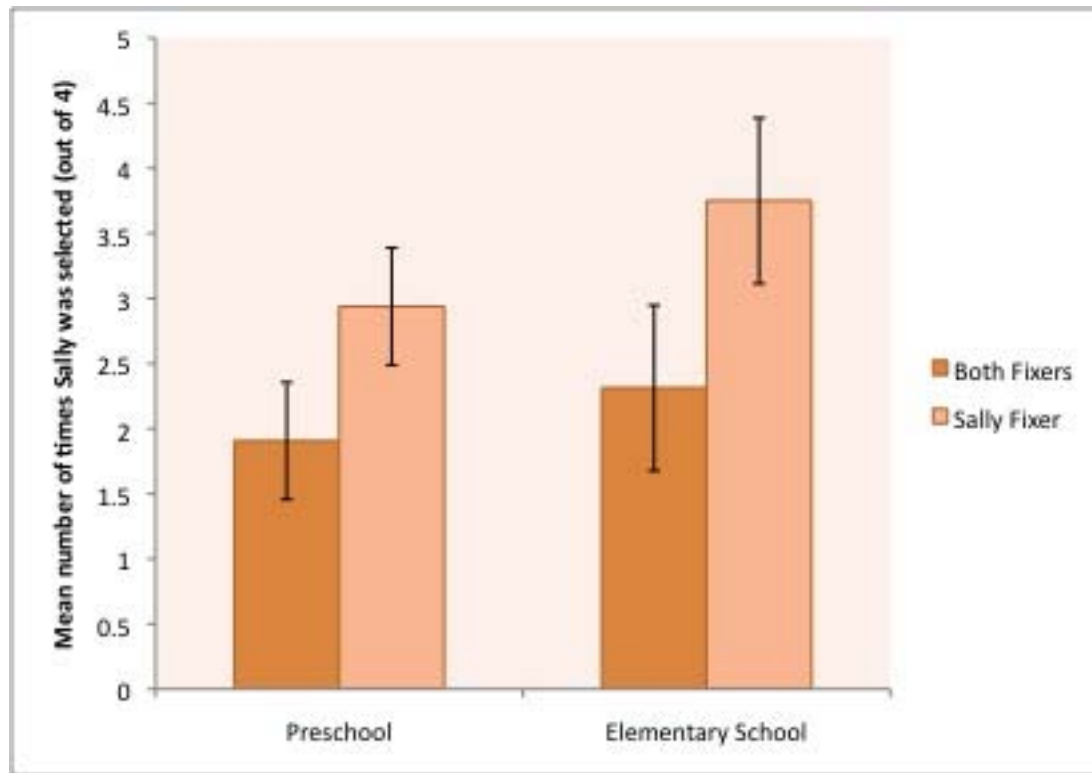
Figure 1. Mean scores of Sally selections by age and condition (out of 4). On the y-axis, 0 = 4

Bill selections for mechanical questions, and 4 = 4 Sally selections for mechanical questions. A

higher score indicates a greater average number of Sally choices. Error bars represent 95%
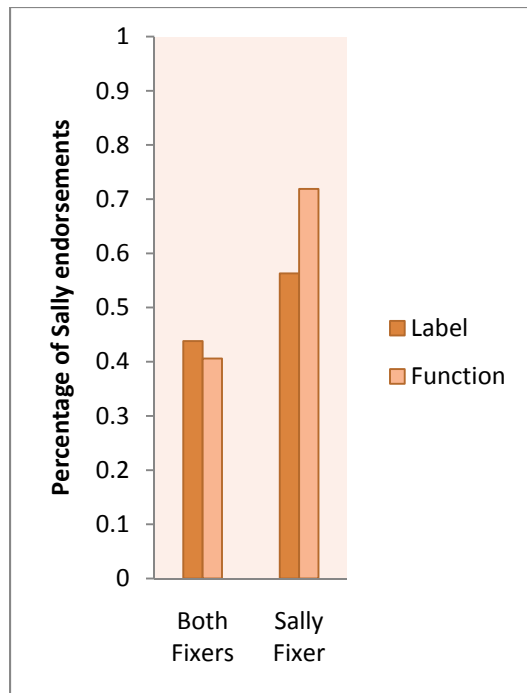
certainty.

Figure 2. Percentage of Sally endorsements for control questions by preschool children by condition (out of 100%). On the y-axis, 1 = 100%. A higher score equals a greater number of Sally choices.
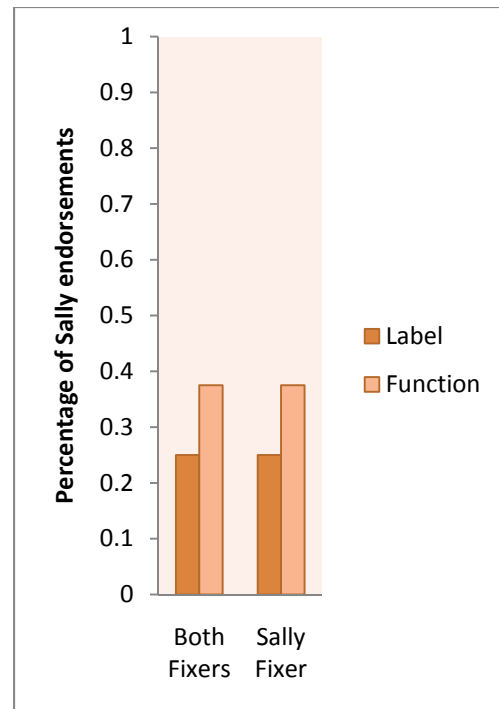

Figure 3. Percentage of Sally endorsements for control questions by elementary school children by condition (out of 100%). On the y-axis, 1 = 100%. A higher score equals a greater number of Sally choices.
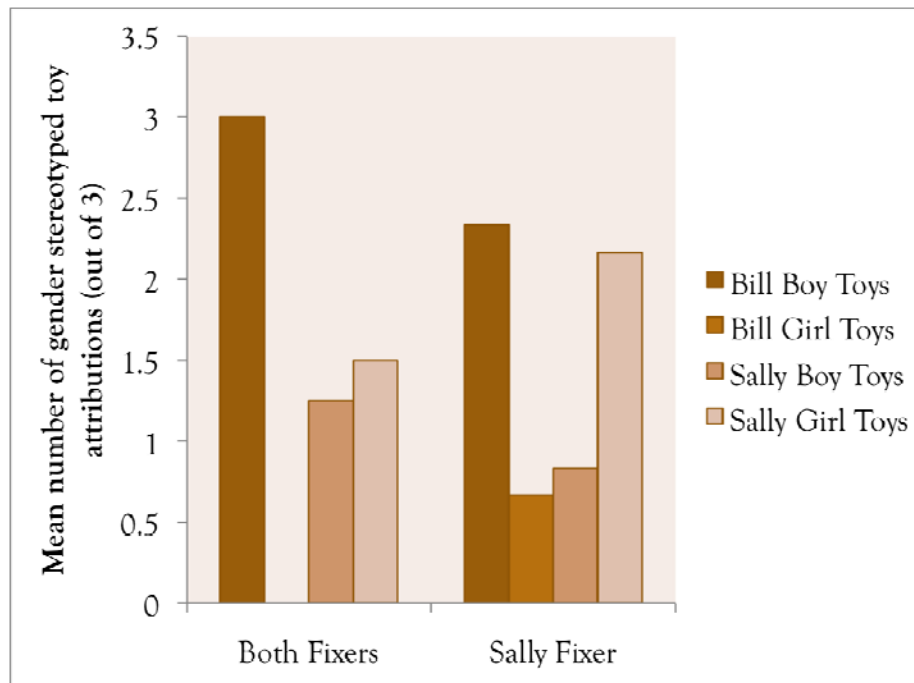
Figure 4: Preschooler's gendered toy attributions to each informant by condition (out of 3).
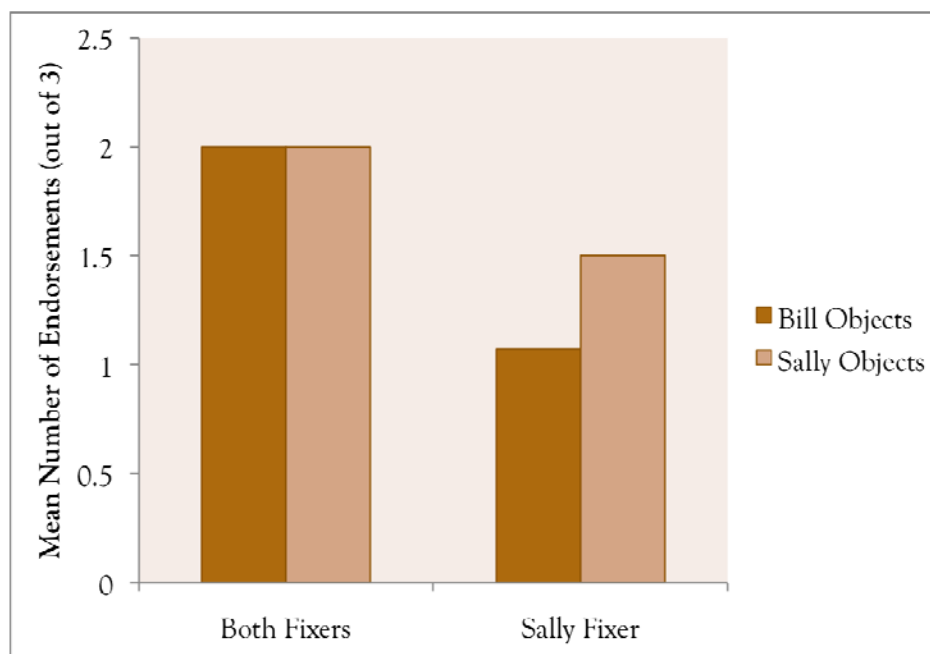

Figure 5. Elementary school children's real world object endorsements for each informant by condition (out of 3).