Lie to Me

Compliant False Accusations by Children

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

by

Amelia Courtney Hritz

August 2014

CHAPTER 1

INTRODUCTION

In February 2014, Dylan Farrow wrote about abuse she suffered at the hands of her father, Woody Allen, in a letter to the New York Times (Farrow, 2014). Farrow described a time when they were playing with trains in the attic and Allen sexually assaulted her. In addition to that incident, she stated that she often did not like the way he would try to be alone with her and stick his thumb in her mouth or be in bed with her wearing only his underwear.

In response, Woody Allen also wrote to the New York Times and categorically denied all abuse (Allen, 2014). Allen described his acrimonious split with Farrow's mother, Mia Farrow, and how she created an "atmosphere of fear and hate towards him" (Allen, 2014 quoting their son Moses Farrow). Allen does not believe Dylan Farrow is lying. Instead, she has false memories of being abused as a consequence of years of repeated suggestions by Mia Farrow. Dylan's piece generated over 3,500 comments ranging from support for Dylan ("[y]ou are brave and strong and I support you in speaking out." (Bonnie M., 2014) to disgust with Allen ("I've boycotted his movies for all these years…Mia fought him as well as she could" (Huntley, 2014). A far fewer number of comments supported Allen's presumed innocence: "there is nothing in Dylan's letter or apparently in the investigation to convince anyone that the abuse really happened" (WHM, 2014). Allen's piece likewise generated close to 2,500 comments that were a mix of support "it would be amazing right now if Mia would just admit publicly that she made this all up due to her jealously" (Julie, 2014) and contempt "Allen's article is a self-serving and vicious attempt to mislead the public about the facts of the case" (Sachs, 2014).

This discussion about what happened to Dylan Farrow has brought attention back to the debate over suggested child memories that inspired a vast amount of research on child memory

when the allegations of sexual abuse by Allen were made over twenty years ago. While we cannot know what happened between Allen and Farrow, previous research has shown that it is possible for adults with biases to alter children's memories through repeated suggestion. Adults with *a priori* beliefs about abuse may single-mindedly attempt to gather only confirmatory evidence and to avoid all avenues that may produce contradictory allegations. Biased interviewers will not ask questions that might require alternate explanations. In addition, when provided with an inconsistent or ambiguous disclosure, biased interviewers either ignore the disclosure or interpret it within the framework of their initial hypothesis. Methods such as repeating specific misinformation**,** offering bribes or threats of punishment, or selectively reinforcing children's statements will alert the child to the interviewer's biased belief. This social pressure can cause the child to make statements that are consistent with the belief of the interviewer rather than the child's experiences (for a review, see Bruck et al., 2006).

Even mild forms of suggestion have been found to increase inaccurate reports by children, such as inaccurate descriptions of events by parents (e.g. Poole & Lindsay, 2001), statements that their peers have already told (Principe & Ceci, 2006) or even natural sharing of memories in conversations with parents and peers (Principe & Schindewolf, 2012).

In addition to causing inaccurate reports, suggestive interviews can also cause children to have false beliefs. Ceci and his colleagues hypothesized that biased interviewers may initially cause children to make inaccurate reports out of compliance to social pressure, but children can come to incorporate these inaccurate statements into their memory (Ceci, Kulkofsky, Klemfuss, Sweeney & Bruck, 2007). Children may come to regard the inaccurate statements as true by making source misattributions: confusing the mental images inspired by the suggestive interview

with the images generated during an actual experience (Ceci, Loftus, Leichtman and Bruck (1994).

False memories can cause children to maintain inaccurate reports in later interviews, this appearing consistent over time. Efforts to retrieve accurate memories following the creation of false memories, such has telling children to say they do not know when they are unsure or to only disclose what actually happened to them, will not offset the impact of the false memory (e.g. Poole & Lindsay, 2001; Zargoza et al., 2001). In addition, even efforts to talk children out of their false beliefs can be unsuccessful (e.g. Ceci et al., 1994).

Historically, the research examining the rates of suggestibility has found that across a variety of paradigms, age is the most powerful predictor of suggestibility proneness, with younger children being more suggestible (see Ceci & Bruck, 1993 for review). Other variables, such as intelligence and language skills, have been found to predict suggestibility, although to a lesser extent (Bruck & Melnyk, 2004; Clarke-Stewart, Mallow & Allhusen, 2004). In addition, recent research has revealed that under certain circumstances, older children and adults are more suggestible and vulnerable to false memory (Brainerd, Reyna & Ceci, 2008). These "reverse age trends" are driven by more developed underlying knowledge and strategies (see e.g., Ceci, Papierno, & Kulkofsky, 2007). For example, greater associations among words can lead adults to be more likely than younger children to remember semantically related words that were never said, such as falsely recalling *sugar* was presented rather than *sweet* (see Brainerd et al., 2008 for review). More research is necessary to determine when legal contexts involve connected meaning or age-related differences in scripts or semantic relatedness. Currently, most studies that involve legal context have found that younger children tend to be significantly more suggestible than older ones (Brainerd & Reyna, 2012).

Previous research has focused primarily on repeated suggestive interviews that contaminate children's memory or cause them to comply with interviewers' beliefs. In contrast to this voluminous literature, in the current study we avoided suggestive questioning and simply requested that children to make a false accusation in order to help others. This alluringly straightforward and non-suggestive approach has not been employed in previous research but would seem to have important implications for court cases in which someone implores a child to make an accusation that at the time was understood to be false by the child (as was alleged by Allen, 2014). In the first study, we sought to answer two questions: (1) will children make a false accusation, merely because someone asks them to do so? And (2) will children maintain their false accusations in later interviews? In a pilot study and a follow-up experiment, we hypothesized that children will be more likely to make an accusation in both interviews if they are younger and more suggestible or have lower intelligence.

In a second study, we further explored the influence exerted on children's memories by adults in the first study. More specifically, the second study examined the different effects of children's perception of the accuracy of the adult's suggestion on their memories. We examined the extent to which children will come to incorporate statements that they initially knew were false into their memories. In addition, we widened our age range to further examine age trends. We hypothesized that both adults and children will be more likely to incorporate an adult's false suggestion into their memory when they believe it is true.

CHAPTER 2

EXPERIMENT 1

Method

Participants. Participants were elementary school-age children in Ithaca, NY. In total there were 70 children, ages 4-12 years old (M=8.53, SD=1.51), and 56% were male. 16 of these participants were enrolled in an elementary school after-school program. The remaining 54 were enrolled in 2-week sessions of summer camp at a children's science museum. Five participants were missing some data due to absences (n=2), equipment failure (n=2), or refusal to complete study (n=1).

Procedure. *Pilot Study*. Two clowns performed a magic show for a group of approximately 40 children enrolled in an after-school program. A research assistant interviewed children individually about the magic show 1-5 days later (M=2.39, SD=1.16). First, children were asked to describe everything they remembered about the show. Then they were asked directed questions about the magic wand i.e. "do you remember seeing a magic wand?" Or, "did the magician do any tricks with a wand?" After the participants described all of their memories about the magic wand, they were shown a broken wand and asked to say the magician broke it (even though he did not). Specifically, the interviewer said the following: "We have heard from children in another school that the magician accidentally broke his wand. I need your help because there is only one magic wand left and we want other classes to get to see the magic show. Can you help me so we can make sure he doesn't do this again? Can you tell me the magician accidentally broke his wand?"

Approximately one week later, a new research assistant interviewed children and asked them to describe the magic show without prompts or suggestions.[1] Then the research assistant explained that the first interviewer had made mistakes. She asked the child if the magician accidentally broke his wand. If the child said yes, the interviewer asked if they had seen it with

[1] While we attempted to interview most children a week after the first interview, a few participants were not available until much later. Therefore our data contain some outliers and time delay ranged from 4 to 46 days (M=9.54, SD=9.8).

their own eyes. Children also completed the Video Suggestibility Scale for Children (Scullin & Ceci, 2001), a psychometric instrument that has been validated on this age group and shown to predict report errors.

Main Experiment. Data for the Experiment 1 were collected at the summer camp at a children's museum in Ithaca, NY. In addition, some changes were made to the procedures after the pilot study. Clowns were replaced with chemists because it may have seemed likely to children that the clown broke the wand due to his clumsiness even if the children had not witnessed it. Instead, a chemist performed "magic tricks" with chemistry, such as mixing acid and base to make a rainbow in a test tube.

The vocabulary subset of the Wechsler Abbreviated Scale of Intelligence by Wechsler (1999) was added to examine if a g-loaded measure of children's verbal intelligence predicted their responses to the first and second interview. Finally, the children were not shown a broken wand (or test tube) during the first interview, because in the pilot study the children seemed to take the presence of a broken wand as strong evidence of the magician breaking it. Thus, the children were not provided with any evidence that the chemist broke a test tube during the show that they saw, only the assertion that others had said he did so during a different camp session.

Results

The number of participants making false accusations are presented in Table 1. The participants who saw the magic show and the broken wand were extremely likely to make the requested false accusation: 15 out of 16 children stated that the magician accidentally broke his wand. Most participants (71%) maintained this accusation in the neutral second interview, but no children said they actually saw it happen.

Participants in the chemistry show group, who were not presented with evidence in the form of a broken test tube, were less likely to make false accusations. However, most of these children (63%) still were willing to state that the chemist broke the test tube. One week later during the neutral interview, fewer children (25%) maintained the accusation, and fewer still (9%) said they saw with their own eyes the chemist break it. In addition, of the 23 who made the false accusation during the neutral 2nd interview, 20 maintained this false allegation despite never having been shown a broken test tube and being told the first interviewer made mistakes. On a positive note, 67% of participants accurately did not make the accusation during the neutral interview. Ultimately 7% of the participants maintained the false allegation and claimed to have witnessed it with their own eyes despite very little suggestion during the previous interview.

Table 1:

Total Number of Participants Making False Accusations

	Magician (T1 n=16) (T2 n=14)		Chemist (T1 n=54) (T2 n=53)		Total (T1 n=70) (T2 n=67)	
	No	Yes	No	Yes	No	Yes
Interview 1						
Nonsuggested Memory	16 (100%)	0 (0%)	54 (100%)	0 (0%)	70 (100%)	0 (0%)
Accusation	1 (6%)	15 (94%)	20 (37%)	34 (63%)	21 (30%)	49 (70%)
Interview 2						
Accusation	4 (29%)	10 (71%)	40 (75%)	13 (25%)	44 (66%)	23 (33%)
Memory	14 (100%)	0 (0%)	48 (91%)	5 (9%)	62 (93%)	5 (7%)

The data from the two experiments were combined and analyzed using logistic regression The effect of age was marginally significant in predicting false allegations during the first interview and maintaining them during the second interview, X^2 (1, N=53)=3.66, p=.06. Six and

seven-year-olds were more likely to comply during the first interview, but were equally likely to maintain the false allegation during the second interview. Interestingly, the 9 participants who stated that they actually saw it happen were all from the youngest age group.

T tests and categorical analysis examining differences in suggestibility and vocabulary among children who made false allegations did not lead to significant results. These tests were underpowered, with only 8% power to detect suggestibility differences and 6% power to detect vocabulary differences at the .05 significance level given sample and effect sizes. Follow-up work with larger and more cognitively diverse (in terms of verbal intelligence and suggestibility-proneness) samples is needed to provide a more robust test.

Individual Differences. *Vocabulary*. Raw scores on the vocabulary subset of the Wechsler Abbreviated Scale of Intelligence were calculated based on the general scoring principles (Wechsler, 1999). These raw scores were converted to T-scores, which are age-corrected and have a wide range of scores (Wechlser, 1999). Results indicated that vocabulary was normally distributed (M=63.53, SD=11.44, n=53) with negative skew. Data is only available for participants who observed the chemistry show. Results indicated that vocabulary was not significantly different for participants who made allegations during T1 or T2. In addition, vocabulary was not correlated with age or suggestibility and was not significantly different for males and females.

Suggestibility. Scores on the Video Suggestibility Scale for Children were normally distributed with a slight positive skew (M=39.08, SD=9.44, n=66). Results indicated that vocabulary was not significantly different for participants who made false allegations during T1 or T2. Suggestibility was not correlated with age or vocabulary and was not significantly different for males and females (all *p's*>.05). In addition, participants who saw the magic show

and participants who saw the chemistry show did not significantly differ on their suggestibility scores. These results were not different when the "yield" and "shift" scores were examined separately, and therefore the combined scores were used for the remaining analyses.

Predicting Which Participants Will Make a False Accusation. We estimated a logistic regression model to predict which participants will make a false accusation in the first interview from show type (magic or chemistry), time delay since show, age, gender, and suggestibility score (from Video Suggestibility Scale for Children). We used dummy codes for all categorical predictors, mean-centered age and standardized suggestibility.

Results of the regression are displayed in Table 2. The probability of making an accusation is 15.42 times the probability of not making an accusation when the show type is magician, time delay is average (2.39 days), age is average (8.53 years), gender is female, suggestibility is average (9.64). None of these predictors significantly changed the odds of making a false accusation. The type of show was marginally significant ($p<0.10$), with the odds of making an accusation in the chemist condition.15 times the odds of making an accusation in the magician condition. When analyses were run without taking into account suggestibility, show type was a significant predictor.[2]

Table 2:

Logistic Regression Predicting who will Make a False Accusation

	B	SE	Odds
Intercept	2.74**	1.14	15.42
Show Type	-1.91+	1.15	0.15
Time Delay	0.10	0.30	1.10
Age	-0.09	0.19	0.92
Gender	-0.50	0.59	0.60

[2] Vocabulary was not measured for the children who saw the magic show. Therefore when analyses were run with the standardized vocabulary variable included as a predictor, data only included the participants who saw the chemistry show (n=54). These results indicated that vocabulary was not a significant predictor.

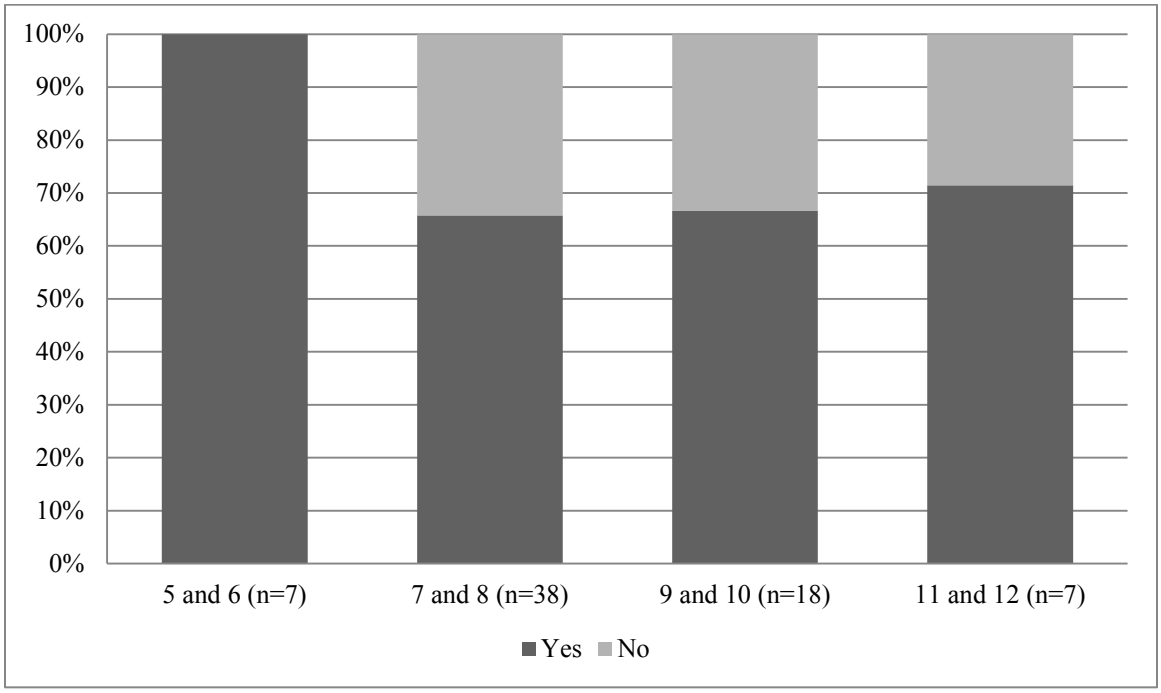
Suggestibility	0.08	0.29	1.08

Notes: ***p*<0.01, **p*<0.05 +*p*<0.10

While age was not a significant predictor it is interesting to note that all seven of the youngest children made the false allegation (see Figure 1). It is possible that with more participants, age would be a significant predictor.

Figure 1:

False Accusations by Age



Predicting Which Participants Maintain the False Accusation. Next, logistic regressions were run to predict who would continue to maintain the false accusation during the second interview. In addition to the variables considered in the previous regression (show type, time delay show, age, gender, suggestibility), this regression took into account whether the participant made an accusation during the first interview. As displayed in Table 3, results indicated that during the second interview, the odds of making the false accusation in the magician condition are approximately 33 times higher than the odds of making an accusation in

the chemist condition ($p<0.01$). During the second interview, with each increase in age of 1 year, the overall odds of making a false allegation multiply by .61, which is statistically significant ($p<.05$). In addition, as the time delay since the show increases by 1 day, the odds of making an accusation multiply by .91, which is marginally significant $p<.07$. Surprisingly, participants' first interview response did not significantly predict their second interview response when all other variables were held constant. Gender and suggestibility continued to not be significant predictors.

Table 3:

Logistic Regression Predicting who will Maintain the Accusation

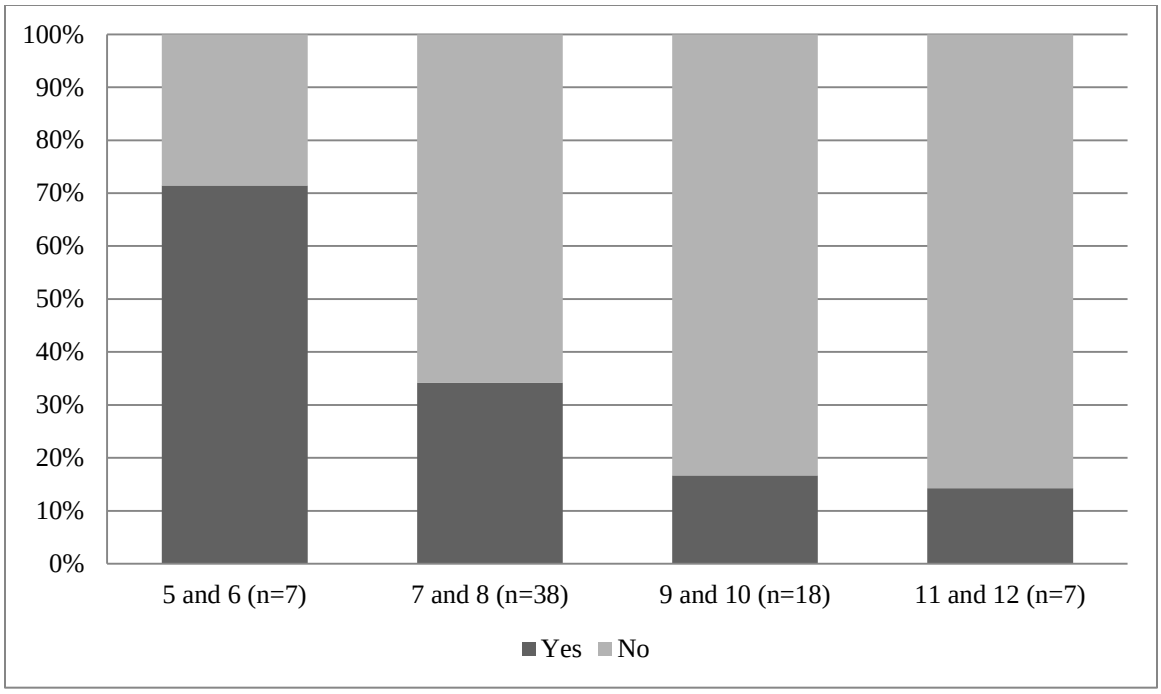
	B	SE	Odds
Intercept	1.53	1.38	4.60
T1 Compliance	0.96	0.76	2.60
Show Type	-3.53**	1.28	0.03
Time Delay	-0.09+	0.05	0.91
Age	-0.49*	0.24	0.61
Gender	-0.19	0.66	0.82
Suggestibility	-0.15	0.34	0.86

Notes: $^{**}p<0.01$, $^{*}p<0.05$ $+p<0.10$

Figure 2 illustrates the age trends in making a false accusation, with the percent of children making false accusations decreasing with age.

Figure 2:

Maintained False Accusations by Age

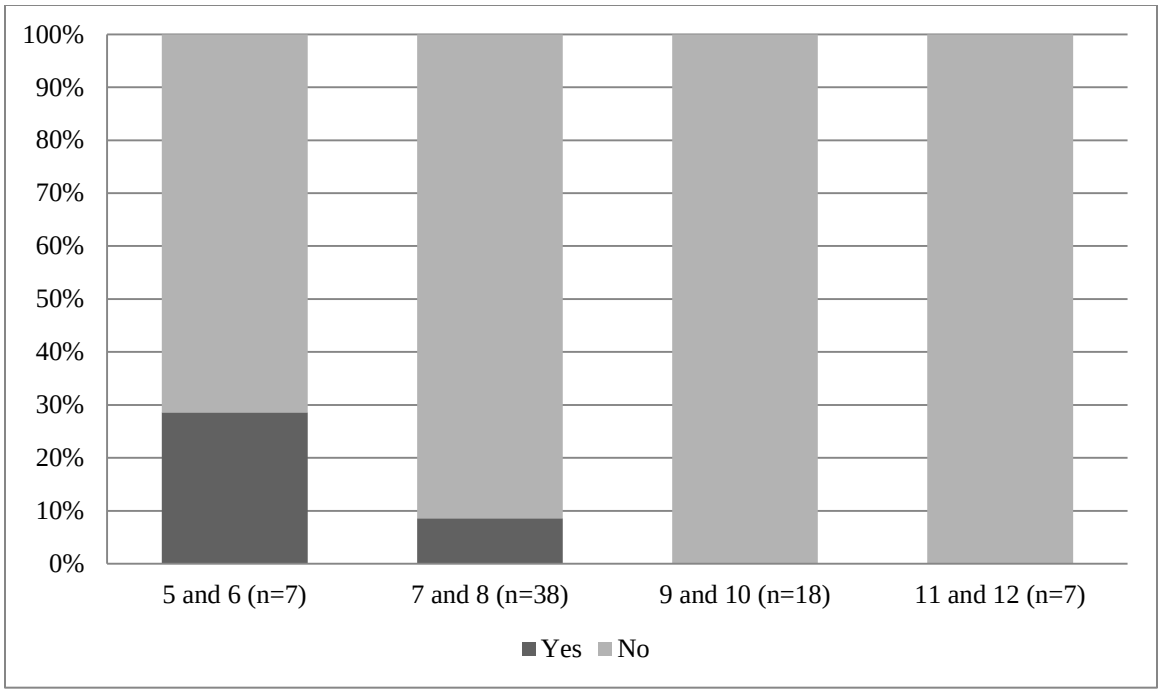


When the regression included vocabulary, which was available only for the participants who saw the chemistry show, the odds of making a false accusation are very low (.01, $p<0.05$) this suggests that after controlling for vocabulary the odds of making an accusation decrease. In addition, the time delay is a significant predictor of maintaining the accusation over time ($p<0.05$), with each additional day increasing the likelihood that and individual with maintain the accusation. Age was a moderately significant predictor of maintaining the accusation, with the probability decreasing as age increases (p=.07). Other predictors were not significant (including vocabulary).

Predicting which Participants Will Say they Saw it Happen. In total, 5 of the children who watched the chemistry show said they saw the chemist break the test tube. All of these children were 8 years old and younger, an age difference that was significant $t(51)=2.02$, $p<0.05$. Figure 3 displays this finding graphically.

Figure 3:

False Memories by Age



CHAPTER 3

EXPERIMENT 2

The pilot study and Experiment 1 found that some children will offer statements in support of requests by a stranger to claim they witnessed an event that they could have. Questions remained about what caused some of the children to incorporate the suggestion into their memory. For example, we wondered if the act of making the false accusation caused participants to maintain it. Alternatively, it is possible that the participants were convinced of the truth of the accusation by the first interviewer. Experiment 2 was conducted in order to explore this finding further. In particular, in the following experiment we manipulated whether the participant knew the suggestion was true of false and measured if this affected the rate at which the participant came to believe the suggestion.

Method

Participants. The second experiment was conducted in a children's science museum as part of an interactive exhibit to teach visitors about human memory. Visitors of all ages who

passed the table could participate in the study, with 297 participants signing up for it, but 28 of them dropped out before completing the final survey. (One participant completed the study twice, so the data from the second survey was dropped.) The participants who dropped out were significantly younger, $t(185.38)=7.21$, $p<.001$, and more likely to be male, $t(291)=-2.02$, $p<.05$.

The remaining 268 participants were between the ages of 3 and 63 ($M=11.79$, $SD=10.38$) and 38% were male, likely reflecting the demographics of daytime museum visitors.

Procedure. Visitors at the children's museum were asked if they would like to participate in our study and learn about memory. Parents or caretakers signed an informed consent document for themselves and/or their children. They also were asked to fill out a demographic information sheet before participating.

Participants were told they would be playing a game called "two truths and a lie" with two research assistants, one of whom would be on their team. The research assistant on their team (RA 1) told participants to look carefully at a picture for fifteen seconds, because they would be playing the game about the picture. Then RA1 put the picture away and explained to the participant that she had come up with two true statements and one false statement about the picture (in reality the second statement is false, but RA 1 told the participant asserted that it is true). The three statements were: a) one of the windows was round, b) there was a newspaper in the driveway of the house, and c) there was a bench in front of the house. If the participant mentioned that the second true statement was actually false (i.e. there was no newspaper in the driveway), the experimenter would repeat that it is true, for example she would say: "I don't think so. I saw a newspaper in the driveway." If the child continued to say that the statement was false, the experimenter would let the child come up with their own true statement and made a note of the change. There were three versions of the house picture that only differed in which

statement was true, so the statement that was actually true systematically varied across conditions as a control for memorability of the three scenes.

RA 1 explained that it will be the participant's role to tell the statements to a second experimenter (RA 2), who will guess which statement is false. RA 1 stressed to participants that in order to win the game (and win stickers) the participant needed to pretend that all the statements were true so it would make it difficult for RA 2 to guess which of the three is false. RA 2 always guessed that the first statement is false (which is always actually a true statement) and so the participants always won. After playing the game, a third experimenter (RA 3) asked the participants true or false questions about the picture, including questions about the statements used in the game. This was done in the absence of RA 1 and RA 2.

Before asking the participant the questions, RA 3 stressed the importance of focusing on the picture instead of the game. RA 3 said "Now, can you think really hard about the picture of the house? Sometimes adults make mistakes, think back to before you played the game and remember what was really in the picture." In addition, RA 3 told participants that if they get a lot of the questions right they will win a second sticker. (In reality, they were always given a second sticker.) Finally, RA 3 asked participants if they noticed any mistakes that RA 1 made when they were playing the game. Afterwards, RA 3 revealed that the second true statement was actually false and gave a brief lesson about memory and suggestibility.[3]

Results

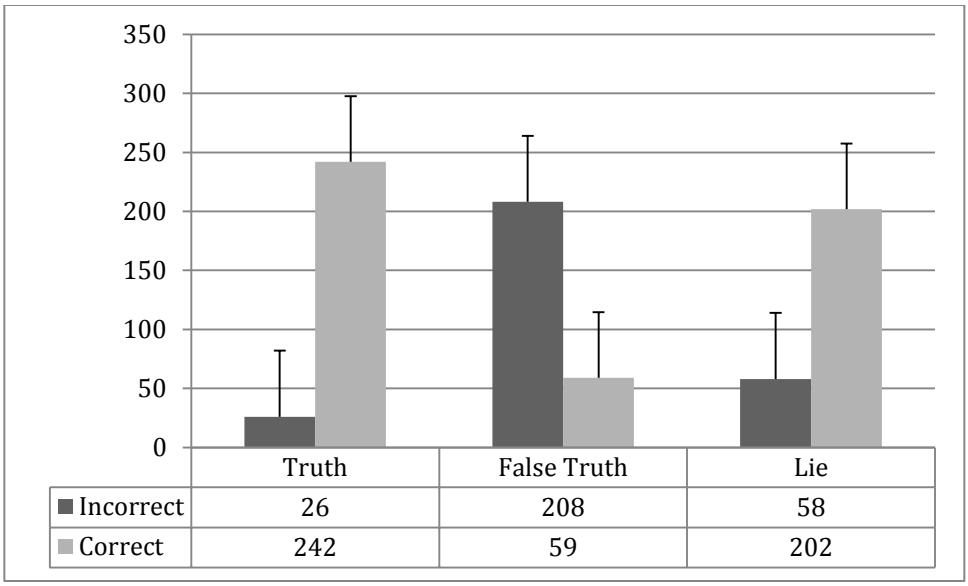
[3] Participants occasionally wanted to make up their own false statements for the game (n=8). These participants are considered to have missing data because we did not ask them about their memory for their own true and false statements on the survey. In addition, one participant correctly identified that the false "true" statement was false and thus created a unique second true statement.

Question Type and Order Effects. Participants were always given the same three statements when they played the game: one of the windows was round, there was a newspaper in the driveway and there was a bench in front of the house. Which of these statements were true, false but described as true, or false varied across 6 conditions. Participants' memory of the true, false-but-asserted-to-be-true, and false statements did not differ based on which statement they were given, $F(1,235)=.04$, $p>.05$, $F(1,235)=.05$, $p>.05$, $F(1,235)=.21$, $p>.05$ (respectively). In addition, participants were given one of three surveys with the order of the questions regarding the target statements varied. The order that participants were asked about the true, false true and false statements on the survey did not significantly change their memories of these statements, $F(1,235)=.96$, $p>.05$, $F(1,235)=3.32$, $p>.05$, $F(1,235)=2.78$, $p>.05$ (respectively). Therefore for the remaining analyses participants are collapsed across condition and survey order.

Memory by Statement Type. Results indicated that most participants answered the question about the false "true" statement (i.e., false-but-asserted-to-be-true) incorrectly on the survey: only 22% correctly said that this statement was false. In contrast, participants were highly accurate on the true and false statements (they were 90% and 78% accurate, respectively). Figure 4 illustrates the number of correct and incorrect responses by statement type.

Figure 4:

Number of Correct and Incorrect Responses by Statement Type



During the final interview, 80-90% of participants' across all ages accurately remembered the true and false "true" statements, $t(259)=-.16$, $p>.05$ and $t(258)=-1.02$, $p>.05$ (respectively). This lack of age effect is illustrated in Figures 5 and 6.

Figure 5:

Percentage of Participants who Remembered the True Statement

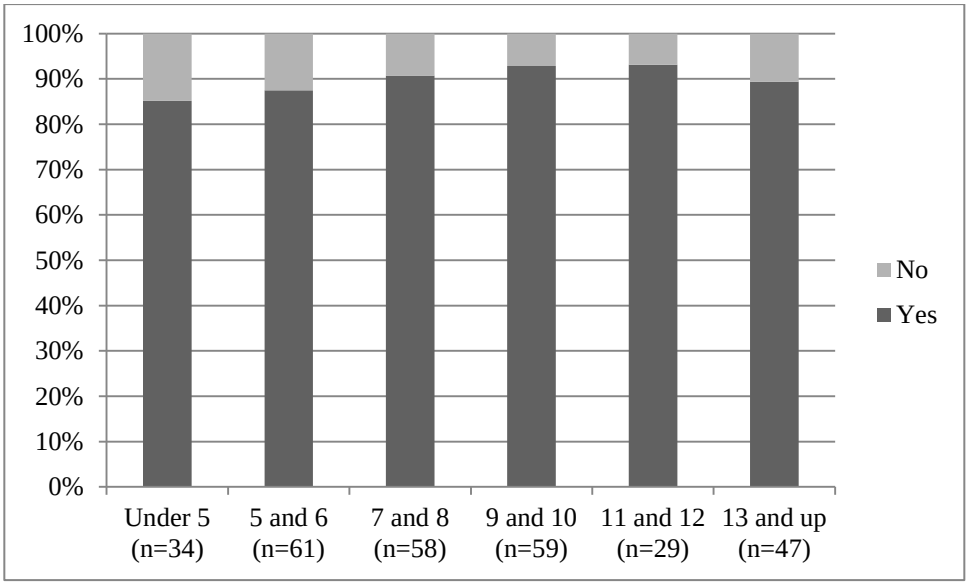
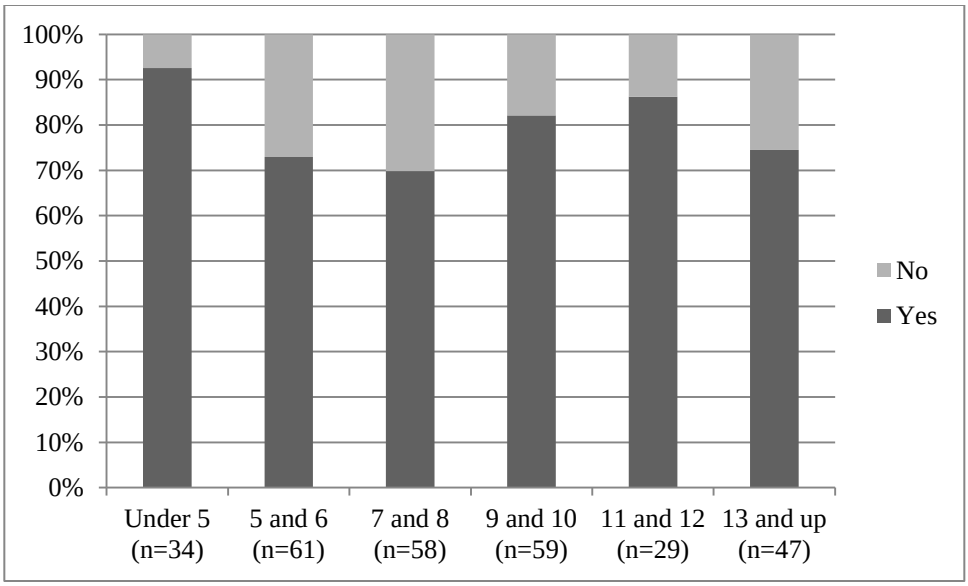


Figure 6:

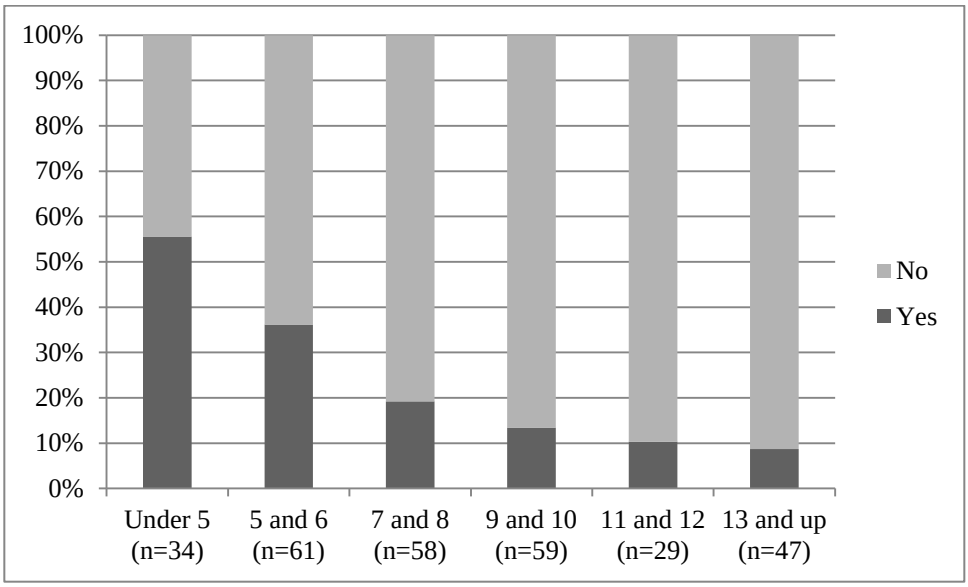
Percentage of Participants who Remembered the False "True" Statement



Age trends did appear, however, when examining only the participants who remembered the false statement as true (see Figure 7); they were much younger (M=7.55) than participants who identified the false statement as false on the survey (M=13.05), this difference is highly significant $t(251)=-3.54$, $p<.001$.

Figure 7:

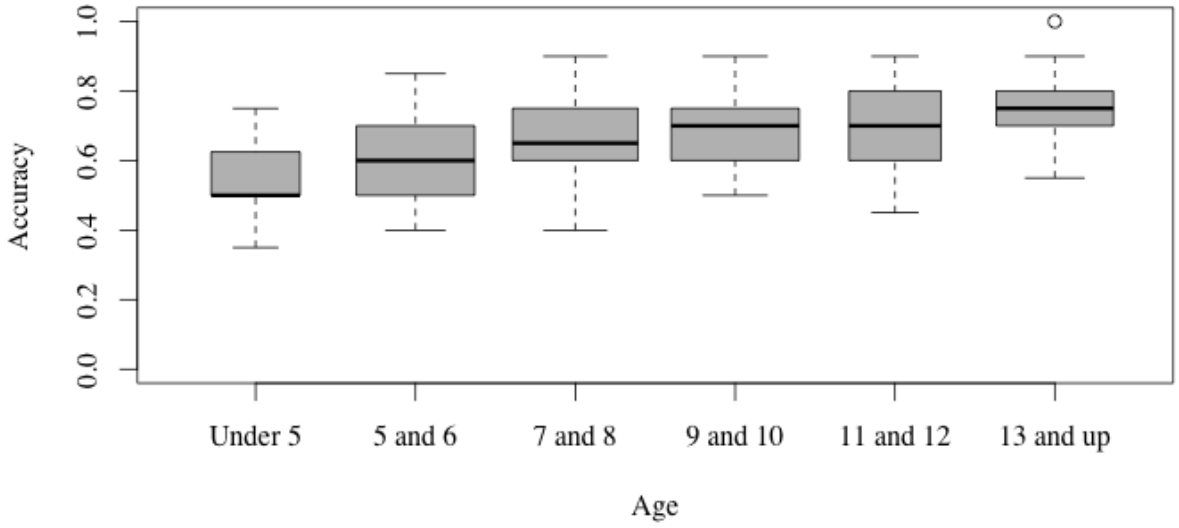
Percentage of Participants who Remembered the False Statement



Survey Accuracy. Participant's overall accuracy on the survey was normally distributed, with scores ranging from 35% correct to 100% correct (M=67%, SD=12%). Participants' overall accuracy on the survey was significantly different if they remembered the false statement $F(265)=-3.62$, $p<.001$ and if they remembered the false "true" statement $F(258)=-7.27$, $p<.001$. Accuracy did not differ based on memory of the true statement. In addition, accuracy was highly correlated with age, $t(259)=7.48$, $p<.001$, with progressively greater accuracy at older ages. Figure 8 displays the first and third quartiles (grey box), mean (dark line), variance (whiskers), and outliers (circles) of accuracy scores by age group. The widths of the boxes are proportional to the square root of the sample sizes.

Figure 8

Boxplot of Accuracy Scores by Age



Individual Survey Item Analysis. In addition to examination of the three survey questions about the true and false statements from the game, the other 17 items on the survey were examined for accuracy and age trends. Results are displayed in Table 4. T-tests revealed

significant differences in average age based on response to multiple items. When there were significant age differences, average ages were higher for correct responses.

Table 4:

Difficulty and Age Differences on Survey Items

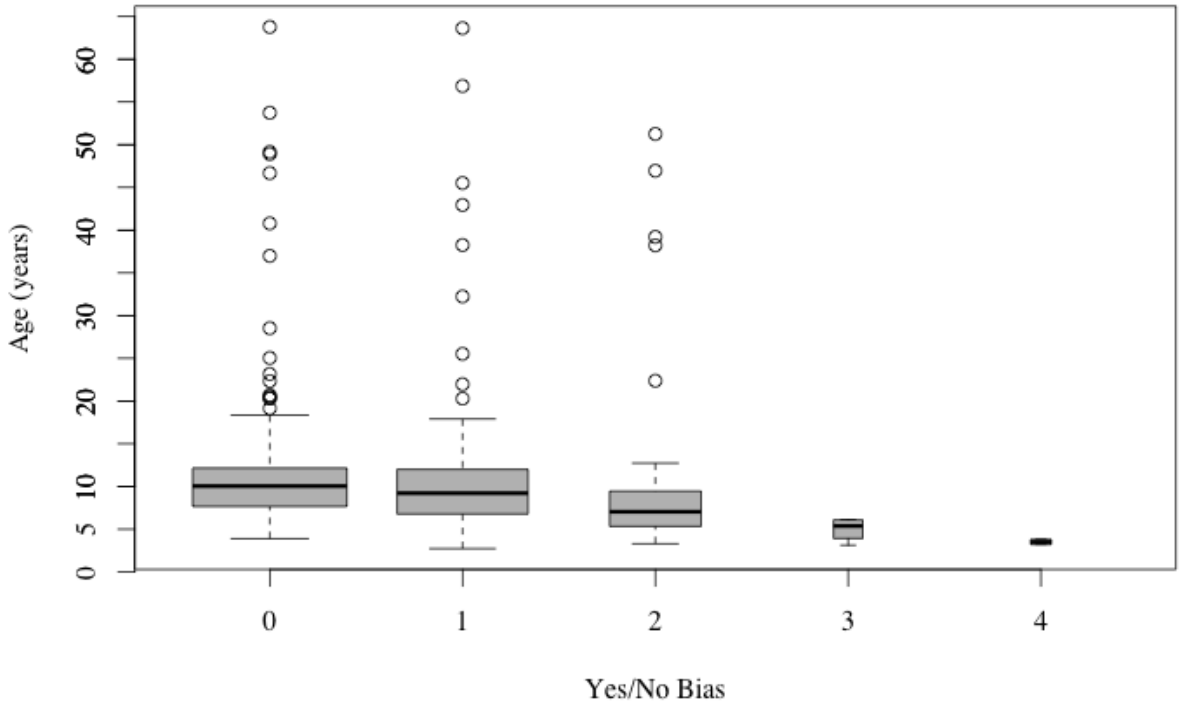
Survey Item	% Correct	Age Differences
Was the house white?	83%	No
Was there a mailbox?	51%	Correct are older
Were there white curtains in the windows?	52%	No
Were there any purple flowers?	79%	No
Were there plants by the street?	57%	Correct are older
Was there a dog in the back yard?	93%	Correct are older
Was there a fence?	69%	Correct are older
Was the front door red?	55%	No
Was there a car in the driveway?	77%	Correct are older
Were there two garage doors?	48%	Correct are older
Was there a swing in the yard?	88%	Correct are older
Was there a tree that was shorter than the house?	82%	Correct are older
Was there a girl waving in one of the windows?	93%	No
Were there black shutters on any of the windows?	31%	No
Was there a light on the garage door?	38%	No
Was there a birdhouse?	75%	Correct are older
Was the roof black?	67%	No

Yes/No Bias. The survey contained 20 questions, of which 10 were true. Participants who answered all "yes" or "no" would thus have an accuracy score of 50%. In order to examine whether certain participants had a yes/no bias, the number of "yes" responses to the survey was examined. The number of yes responses ranged from 2 to 20 (M=8.77, SD=2.62). A yes/no bias score was calculated by taking the absolute value of the z-score of the number of yes responses. For example, if yes/no bias = 1, the absolute value of the participant's number of yes responses was within 1 standard deviation from the mean number of yes responses. Results indicated that as the deviation from the mean of yes/no bias increased, the average age of participants decreased significantly $F(1,259)=4.25$, $p<.05$.

Figure 9 displays the first and third quartiles (grey box), mean (dark line), variance (whiskers), and outliers (circles) of age in years by yes/no bias. The widths of the boxes are proportional to the square root of the sample sizes. There are very few participants with yes/no bias scores of 3 and 4 (n=4 and n=2, respectively).

Figure 9:

Boxplot of Age and Ye/No Bias



Predicting Erroneous Memory of False "True" Statement. We estimated a logistic regression model to predict which participants would remember the false "true" statement even though the item was not in the picture. We used a number of variables that we hypothesized may be linked to a false memory (yes/no bias, age, gender, and memory of the true and false statements). In this model we used dummy codes for all categorical predictors, mean-centered

age and standardized accuracy and yes/no bias. The estimates of the raw scores of the predictor variables on memory of the false "true" statement, standard errors and odds ratios are displayed in Table 5. The positive estimates indicate that memory of the false "truth" is more likely.

Results of the regression indicated that the probability of erroneously remembering the false "true" statement as true is 0.16 times the probability of not remembering it when accuracy is average (.67), yes/no bias is average (8.77), age is average (11.79), gender is female and inaccurate responses are given for the true or false statements. Accuracy on the survey significantly predicted memory of the false "truth," with the odds of having a false memory multiplying by 0.37 when accuracy increases by 1 standard deviation. In addition participants inaccurate memories of the true and false statements significantly increased their odds of incorrectly remembering the false "true" statement.[4]

Table 5:

Logistic Regression Predicting Memory of False Truth

	B	SE	Odds
Intercept	-1.84**	0.64	0.16
Age	0.01	0.02	1.01
Gender	0.35	0.35	1.42
Accuracy	-0.99***	0.23	0.37
Yes/no bias	0.27	0.29	1.30
No memory of truth	2.00***	0.49	7.42
Memory of lie	1.58***	0.44	4.86

Notes: ***p<0.001, **p<0.01, *p<0.05

Predicting Memory of False Statement. In this next analysis we examined participants' memories for the statement that was identified by the research assistant as false during the game. We estimated a logistic regression model to predict which participants would accurately remember that the false statement was false using the same variables (accuracy, yes/no bias, age,

[4] When memory of the false "true" statement is estimated with only age and gender entered as predictors, neither variable significantly changed the odds of a false memory.

gender, and memory of the true and false "true" statements). We again used dummy codes for all categorical predictors, mean-centered age and standardized accuracy and yes/no bias.

Results of the regression are displayed in Table 6. In Model 1, when only age and gender are entered into the model, age is a highly significant predictor of which participants will erroneously remember the false statement as true, with younger participants being more likely to remember the false statement. In Model 2, when all predicting variables are included, age is no longer a significant predictor of memory of the false statement. Accuracy, yes/no bias and memory of the false "truth" are all significant predictors of memory of the false statement. While all other variables are held constant, as accuracy decreases, the probability of remembering the false statement increases. Conversely, as yes/no bias increases, the probability of remembering the false statement also increases. Finally, participants who remember the false "true" statement are significantly more likely to remember the false statement. These findings suggest that lower accuracy and increased yes/no bias, which were highly correlated with age, explain the effect of age that we found in the first model.

Table 6:

Logistic Regression Predicting Memory of False Statement

	Model 1			Model 2		
	B	SE	Odds	B	SE	Odds
Intercept	-1.80***	0.30	0.16	-3.16***	0.67	0.04
Age	-0.19***	0.05	0.83	-0.05	0.04	0.95
Gender	0.04	0.33	1.04	-0.03	0.36	0.97
Accuracy				-1.13***	0.24	0.32
Yes/No Bias				0.46*	0.23	1.59
Truth Response				0.76	0.57	2.13
False "True" Response				1.63***	0.45	5.10

Notes: ***p<0.001, **p<0.01, *p<0.05

Taken together, these findings paint a mixed picture of developmental differences in accepting false statements. At first it seems as if memory of the false statement is predicted by

age, but upon closer examination it appears that the false memory is better predicted by weaker memory for the picture and a bias towards giving all "yes" or "no" answers.

CHAPTER 4

DISCUSSION

Results of the first experiment demonstrated that children would comply with an adult's request to make an accusation even if they do not have first-hand knowledge of its accuracy. On a positive note, results also indicated that most children would accurately disclose the truth in a neutral interview. The significant differences in results in the chemistry show and the magic show illustrate how contextualized these results are. With more evidence (brandishing a broken test tube) or overt use of stereotypes (such as calling the chemist clumsy), it is likely that we could have further increased the rates of compliance. Interestingly, the initial compliance may have created false memories in some of the youngest children because they subsequently claimed to remember seeing the false event despite being given "release" instructions by the neutral interviewer.

The first experiment was limited in that we were not able to induce the stress of a real interview in which children a) knew the person they were accusing, and b) understood that this individual could face adverse consequences as a result of their false statement. Ethical considerations required that we did not make the children feel uncomfortable and thus we told children that it was probably an accident, but we needed their help to make sure this did not happen again. Even though there were no negative consequences to making the false accusation, it is likely that there was also less pressure exerted on participants to make a false accusation. Analyses of real interviews with children who were suspected victims of being abused while in daycare revealed that multiple repeated suggestive methods were used such as introduction of

new suggestive information, positive reinforcement, disbelief, conformity pressure and invitations to pretend (Schreiber et al., 2006). Thus, in real cases, interviewers with biases may exert more pressure and use more suggestions than was done in the current study.

After the first experiment, questions remained about why the children maintained the false accusation and what created the seemingly false memories in the younger children. Was it the act of making the false accusation or the incorporation of the adult's suggestion? Anecdotally, children seemed to believe the interviewer, and did not feel they were lying when they made their false accusations. In order to answer these questions, in the second experiment we showed participants a picture and provided them with statements about the picture that we said were true and false and asked them to play a game in which they try to trick another person into believing that they were all true. After warning participants to think about the picture instead of the game, and testing their memories for the picture as opposed to what was stated during the game, results indicated that both children and adults were likely to incorporate false statements into their memory when they believed they were true. In addition, younger participants also incorporated the statements into their memory that they had previously known were false. Upon closer inspection, the significant age trends in remembering the false statement could be explained by weaker memory accuracy, biases towards all "yes" or all "no" answers, or willingness to believe the false "true" statement.

Upon closer inspection of the individual items in the survey, significant age trends appeared for a number of items. All of the significant age trends revealed that older participants were more accurate. These findings seem to conflict with the research on reverse age trends that suggests that adults are more likely to have spontaneous false memories of thematically connected items (see Brainerd & Reyna, 2012, for a review). Based on this research one might

expect adults to be more likely to falsely remember a mailbox in the picture because a mailbox is associated with a house. Our results may be due to the fact that children also have strong associations between such items related to house. It is possible that if the picture were about something less familiar to children, such as an office, reverse age trends would appear. However, reverse age trends is but one of myriad forces at play in producing memory inaccuracies. Other forces include mainly factors that protect older individuals from making memory errors, such as stronger traces and greater monitoring. Any full accounting of age differences in remembering will need to take into account all forces, not just associative strength of items. It is possible that even when older individuals possess stronger associations that place them at greater risk of making autosuggestions, they are on net less likely to do so because their stronger memory traces and more effective monitoring buffer them from doing so.

Overall, the findings of experiments 1 and 2 have implications for cases when adults with biases provide evidence or other clues during conversations with suspected victims. It seems possible that this can cause some young children or individuals with weak memories to make accusations that they know, at least initially, are false. In addition, in later interviews some of these individuals may come to believe that their initial false report is true. Furthermore, if the interviewer is persuasive enough to convince children and adults that their false belief is accurate, both adults and children may come to have memories that are consistent with the interviewers beliefs rather than with their own experiences.

References

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: a review of data and theory. *Psychological bulletin*, *134*(3), 343.

Bruck, M. & Ceci, S. J. (2006). Children's suggestibility: Characteristics and mechanisms. *Advances in Child Development and Behavior*, *34,* 247-281.

Bruck, M. & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology, 50,* 419-428.

Bruck, M., Ceci, S. J. & Principe, G. (2006). The child and the law. In I. Sigel & A. Renniger (Eds.), *Handbook of child psychology: Child psychology in practice* (6th ed., Vol. 4, pp. 776-816). New York, NY: Wiley.

Bruck, M., & Melnyk, L. (2004). Individual differences in children's suggestibility: A review and synthesis. *Applied Cognitive Psychology*, *18*(8), 947-996.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony.* Washington, DC: American Psychological Association.

Ceci, S. J. & Bruck, M. (2006). Children's suggestibility: characteristics and mechanisms. In: Robert V. Kail, Editor(s), *Advances in Child Development and Behavior*, *34*, 247-281.

Ceci, S. J., Loftus, E. W., Leichtman, M., & Bruck, M. (1994). The role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis, 62,* 304-320.

Ceci, S. J., Papierno, P. B., & Kulkofsky, S. (2007). Representational constraints on children's suggestibility. *Psychological Science*, *18*(6), 503-509.

Clarke‐Stewart, K. A., Malloy, L. C., & Allhusen, V. D. (2004). Verbal ability, self‐control, and close relationships with parents protect children against misleading suggestions. *Applied Cognitive Psychology*, *18*(8), 1037-1058.

Dunn, L. M., & Dunn, L. M. (1997). *Examiner's manual for the PPVT-III Peabody picture vocabulary test*. AGS.

False accusations in an investigative context: Differences between suggestible and non-suggestible witnesses

Poole, D. A. & Lindsay, D. S. (2001). Children's eyewitness reports after exposure to misinformation from parents. *Journal of Experimental Psychology: Applied, 7,* 27-50.

Principe, G. F. & Ceci, S. J. (2002). "I saw it with my own ears": The effects of peer conversations on preschoolers' reports of nonexperienced events. *Journal of Experimental Child Psychology, 83,* 1-25.

Principe, G. & Schindewolf, E. (2012). Natural conversations as a source of false memories in children: Implications for the testimony of young witnesses. *Developmental Review, 32,* 205-223.

Schreiber, N., Bellah, L. D., Martinez, Y., McLaurin, K. A., Strok, R., Garven, S. & Wood, J. M. (2006) Suggestive interviewing in the McMartin Preschool and Kelly Michaels daycare abuse cases: A case study. *Psychology Press, 1,* 16-47.

Scullin, M. H. & Ceci, S. J. (2001). A suggestibility scale for children. *Personality and Individual Differences, 30,* 843-856.

Wechsler, D. (1999). *Wechsler Intelligence Scale for Children-IV*. Psychological Corp. San Antonio, TX.

Zargoza, M. S., Payment, K. E., Kichler, J., Stines, L., & Drivdahl, S. (2001). Interviewing witnesses: forced confabulation and confirmatory feedback increase false memories. *Psychological Science, 12,* 473-477.