DID SHE DO IT ON PURPOSE? : YOUNG CHILDREN'S PERFORMANCE ON

YES/NO AND MULTIPLE CHOICE QUESTIONS ABOUT INTENTIONS

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

Jessica Zoe Klemfuss

January 2008

ABSTRACT

While children demonstrate the ability to understand others' intentions as early as 6 months of age, they are not able to verbally answer questions about intentions until between 3 and 6 years of age. The present study was designed to explore whether different question formats influence the accuracy of children's responses to questions about intentions, and whether the development of executive attention plays a role in children's ability to answer questions about intentions. Seventy-nine preschool children between 3 and 5 years of age were recruited for this study. Children viewed videos of playground events involving two main characters. Within this video there were three events which were performed either on purpose or by accident. After the video, children were asked a series of questions including prompts about the intentions behind the three target events. Children were significantly worse at answering questions about intentions as compared to questions which did not involve intentions. Children demonstrated an overall "yes" bias for intentions and non- intentions questions. For intentions questions, this means that they were most accurate when they were asked yes/no questions where the answer should be "yes" (Match) and least accurate when they were asked yes/no questions where the answer should be "no" (NonMatch). They were equally accurate at Match and forced-choice questions. There was an interaction with age for the non-intentions questions so that younger children were more prone to demonstrate a "yes" bias and show increased accuracy for Match questions and decreased accuracy for NonMatch questions. Younger children were equally accurate at Match and forced-choice questions, and more accurate at forced choice questions than NonMatch questions. Executive attention measures predicted children's accuracy for NonMatch, non intention questions. The results are discussed in terms of forensic applications.

BIOGRAPHICAL SKETCH

Jessica Zoe Klemfuss received her B.A. in Psychology from the University of California at Berkeley in 2004. At U.C. Berkeley she worked in the Cognitive Development lab where she assisted in projects involving children's understanding of adult preferences, and infants' abilities to learn language in differing contexts. She then worked as a research assistant in an fMRI lab at the University of California at San Diego, studying the effects of Obstructive Sleep Apnea before coming to Cornell as a graduate student in the field of Human Development in 2005. Her academic interests include children's eyewitness memory accuracy, interview factors relating to children's testimonial accuracy, and children's abilities to communicate about complex concepts such as intentionality.

To my family.

ACKNOWLEDGMENTS

My gratitude goes first to my advisors, Steve Ceci and Qi Wang for their valued advice and support throughout this process. Steve has gone out of his way to be available from the initial point of discussing this idea through reading the drafts of this thesis. His support has been extremely useful, as has the freedom to explore ideas that interest me. Qi has also been a valuable member of my committee. From my first week at Cornell I have had opportunities to engage in research in her lab and practice the skills I would need to carry out a project of my own. Further, both Steve and Qi's productivity and impeccable research records have been both inspirational and motivating to me.

I was also fortunate to have five very capable undergraduate assistants on this project- Christina Han, Karyn Hartz, Alex Kaplan, Bethany Ojalehto, and Ari Silbermann. Without their help this research would not have been possible. They were all extremely talented and eager and I have no doubt they will each be extremely successful in their own careers.

I would also like to thank my family for their continuing support and patience. This gratitude extends far beyond my graduate experience. It is my parents' combined advice that has kept me loving the academic pursuit while still having a life outside of academia. They have fostered my creativity and made me proud to achieve. My sister Nola has ceaselessly impressed me with her successes, creativity, and intelligence and though she is younger than I am, I admire her achievements, and in some ways, model myself after her.

Finally, I am grateful to David Brown who helped me with everything from data entry to listening to me discuss and present my ideas and data. Most importantly, he has provided unconditional support in every aspect of this project and my career thus far.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Research on the development of intentional understanding has taken two
complementary perspectives. The first line of research is largely cognitive and
emphasizes children's early abilities to understand the intentions of others. The second
line, focusing on linguistic research, has demonstrated the limitations of young
children's understanding of intentions. Studies have shown that although the ability to
comprehend others' intentions may exist at a very young age (Behne, Carpenter, Call,
& Tomasello, 2005; Meltzoff, 1995; Shimizu & Johnson, 2004), children may not be
able to verbally communicate this understanding effectively until much later (Baird &
Moses, 2001; Piaget, 1932; Smith, 1978).

Infants appear adept at understanding others' intentions in non-linguistic
situations. Meltzoff (1995) demonstrated 18 month old children could replicate an
adult's intended action, even when that action was not successfully performed by the
adult. In this study, adults attempted to perform tasks involving manipulating two
objects, and repeatedly failed. When 18-month-olds were given the same objects they
performed the intended action and not the failed action that the adult had performed,
indicating that intentional understanding begins developing well before children
become competent language-users. Since Meltzoff's study, other researchers have
attempted to identify the earliest indications of intentional understanding (Behne et.
al., 2005; Hamlin, Wynn, & Bloom, 2007; Legerstee, Barna, & DiAdamo, 2000;
Moses, Baldwin, Rosicky, & Tidball, 2001), demonstrating intentional competence as
early as 6 months of age (Hamlin et al, 2007; Legerstee et al., 2000).

Early studies examining children's abilities to understand others' intentions
focused on children's limitations in verbally communicating about intentions (Piaget,

1932; Smith, 1978). Piaget (1932) presented young children with images of adults performing intentional and unintentional actions and then questioned the children about whether or not the adults were morally at fault for those actions. Preschool-aged children based their judgments more on the effects of the acts than on the apparent intention behind them, implying that children do not fully understand the influence of intentions on one's behavior until late in their preschool years. Since Piaget's work, other researchers have found conflicting results concerning children's abilities to verbally communicate about intentions. While some data have supported Piaget's findings (Astingtion & Lee, 1991; Baird and Moses, 2001; Schult, 2002; Smith, 1978), other studies have demonstrated more precocious competence in the domain of intentionality (Moses, 1993; Shultz, Wells, & Sarda, 1980). The earliest estimates suggest that children may be able to answer questions about intentions by three years of age (ex: Moses, 1993), while later estimates suggest that children are not consistent at this ability until as late as six years of age (ex: Smith, 1978).

Given the evidence that children can comprehend the intentionality behind actions before their first birthday but do not demonstrate the same competence when language is involved, the challenge for preschool-aged children is likely more linguistic than conceptual. However, children are rarely inhibited from attempting to answer questions when they are unsure of the correct answer (Battin & Ceci, 2003), and more importantly, they may often sound credible when they answer closed-ended questions about intentions (ex: Kulkofsky, Klemfuss, & Ceci, unpublished raw data).

For example, preliminary data from a previous study suggest that children do in fact have difficulty answering directed questions about intentional acts of the form "Did she do it on purpose?" (Kulkofsky et al, 2007). In this study children witnessed a staged event during which "Miss Baker" baked cookies for her friend's birthday in class with the children. One week later the children were given the opportunity to

recall everything they remembered about the event, and then the interviewer asked a series of misleading questions about the event. If children did not assent to these false questions, the questions were followed up with increasingly more suggestive techniques. First with peer pressure "are you sure? The other kids told me she did", next with authority pressure "it's really important I know the truth, did _____ happen", and lastly, with a prompt to imagine that the event did in fact take place and describe it.

The question of interest to the present study is "did Miss Baker pour too much salt on purpose". Over 70% of children in the Kulkofsky et al. study incorrectly answered that yes, she did after just one interview. This result is particularly dramatic given that the other rates of suggestibility were 60.8% for questions suggesting that "Miss Baker" wore a yellow rather than a blue apron, 47.1% for the suggestion that she brought gummy bears rather than M&Ms, 33.3% for questions suggesting that the children had been allowed to eat the cookie dough (they had not), and 21.6% of children assented that Miss Baker had placed a sticker on her knee when in actuality she handed it to them. Interestingly, children were often more accurate when they were immediately asked to elaborate. For example, when one child was asked, "did Miss Baker pour too much salt on purpose", she responded "yes she did. I know that". However, when the interviewer followed up with the question "What can you tell me about that?" the child answered "...So she put some salt in it, then she poured too much, and then that was an accident". Therefore, this child falsely assented to the intentionality question, but then immediately afterwards in free recall, provided a correct response using intentional language correctly. Importantly, in this example, the child used more assertive language in the false assent than in the open-ended question, making her appear more credible when she in fact was less accurate.

While the studies mentioned above illustrate the limitations in children's abilities to accurately answer questions about intentions, and hint at the unfortunate conclusion that inaccuracy for these types of questions can coincide with children sounding credible, they do not sufficiently address the issue of why this happens.

One explanation for children's difficulty in answering questions about intentions is that the language is too complex or that the format of the questions is too difficult to monitor for the intentions of the interviewer. In other words, the children may be unclear as to what the interviewer wants to know. For example, language that adults typically associate with intentionality may be beyond preschool aged children's understanding. Further, questions about intentions may be more complex than questions that do not involve intentions because in essence, they have an extra clause imbedded (ex: Did she do it [on purpose]". Indeed, when interviewers use more complex language in child-directed questions, children's accuracy is lower than when interviewers use simple, age-appropriate language (Carter, Bottoms, & Levine, 1996). Although questions about intentions may be unnecessarily confusing in many situations, in other cases, it may be critical to know the intentions behind an act that only a young child witnessed as in the case of child injury. Distinguishing between accidental harm and intentional harm inflicted on a child would likely have very different legal consequences.

The leading nature of yes/no questions provides a second explanation for children's mistakes when answering questions about intentions. Suggestibility research demonstrates that closed-ended questions as opposed to open-ended questions may lead to increases in false assents and suggestibility (Peterson, Dowden, & Tobin, 1999; Peterson & Bell, 1996; Poole & Lindsay, 1995), particularly when asked by an authority figure (Ceci, Ross, & Toglia, 1987). The complex and confusing nature of questions about intentions may heighten young children's suggestibility. These two

explanations may work together so that when a child is confused about how to answer an intentional question, they may use the credibility and perceived perspective of the interviewer, coupled with interviewer-provided suggestions, to falsely assent to the question.

Further, previous research has indicated that even when children do not have sufficient knowledge or ability to answer an interviewer's question, they try anyway (ex: Brady, Poole, Warren, & Jones, 1999; Peterson & Grant, 2001). This may be because they feel pressure to perform well for the interviewer (Ceci and Bruck 1993). It may also be because children learn about conversational strategy before they learn that responses to questions should be accurate (Steffenson, 1977, cited in Choi, 1991). In other words, children learn that they should provide some sort of response to a yes/no question before they learn that this response should be correct. Further, children may be practicing their understanding of intentional language. Language has been demonstrated to assist in the development of such related concepts as theory of mind by giving children practice talking with other people and by allowing mothers to help scaffold children's understanding (Astington, 1999). Children may be subconsciously attempting to answer these questions in hope of getting positive or corrective feedback. Therefore, perhaps children use the language of intention before they are able to reliably connect the concept of intentionality to the words for intentionality. This trial and error conversation strategy is particularly dangerous in a forensic setting because children often sound more accurate than they are (Battin, 2004).

These results combined with a body of research on children's accuracy indicate that if interviewers rely solely on open-ended questions, children can be highly accurate, even about complex, or confusing topics. However, despite the fact that every forensic interview protocol to date recommends the use of free recall and open-

ended questions, the rates of open-ended question used in practice are dwarfed by the number of closed ended questions. Lamb and his colleagues (1996) found that only 2% of questions used in a sample of forensic interview transcripts were open-ended. Other studies which have examined the rates of yes/no questions in actual interview transcripts estimate their rate of usage in forensic interviews at between 64 and 73 percent (McGough & Warren, 1994; Walker, 1993; Warren et al, 1996).

In addition to the main issue of whether preschoolers are able to accurately answer questions about intentions, it is also interesting to determine why they are making these errors. One potential explanation is that children may misunderstand questions such as "Did she do it on purpose?" because they misinterpret which part of the question the interviewer is interested in (did she do it vs. do it intentionally). In a sense, when children are interviewed about another person's intentions they must keep in mind not only the intentions of the actor in the observed event, but also the intentions of the interviewer. They must understand that the interviewer is seeking information regarding the intentions rather than outcome of an event, the interest of most interviewers.

If, for example, children are more adept at answering questions about actions than mental states, they may be more competent at answering descriptive, action-based parts of a question such as "Did the grown up bump into the littler girl?" as opposed to the intention portion of the question "Did the grown up bump into the little girl on purpose?". The second part of the question, which an adult would recognize as the portion the interviewer actually wants answered, involves understanding both the intentions of the actor and of the interviewer. Children's mistakes in this type of question may occur because to them, questions about intentions are compound questions, and they answer the portion about the act itself ("Did the grown up bump into the little girl?") before they answer about the intentions behind the act. Children

may also answer based on the "action" portion of the question because of underdeveloped inhibitory control (Frye, Zelazo, & Palfai, 1995). Young children may not be able to inhibit their response long enough to wait for the full question. If this were the case, children would have already processed their response to the first part of the question ("Did the grown up bump into the little girl?") before they even begin to think about the second part of it ("Did she do it on purpose?").

In summary, I hypothesize that children's difficulties in answering interview questions about intentions may be a product of leading influences and difficulties in monitoring complex questions. In the present study, I compared children's responses to yes/no questions with those from forced choice questions, the latter of which make salient both parts of the interviewer's request—was the act performed and was it performed deliberately. I propose that yes/no questions are leading about both the act itself, and more importantly, about the intentions of the act because yes/no questions necessarily suggest one option (in this case, on purpose, or by accident). In contrast, forced choice questions imply that the act itself took place, but offer no judgment about the intentions behind the act because they provide all of the possible solutions (again, on purpose, OR by accident). Secondly, forced choice questions require an answer specifically about the intentions of the act because it only allows for two possible answers "on purpose" or "by accident". The forced choice may serve as a cue indicating that the interviewer is more interested in the actor's intentions than in the act itself.

Because I hypothesize that children's errors are often based on underdeveloped monitoring skills, I also administered two executive attention measures to shed light on the process(es) through which children regulate their responses and select the key components of the interview questions. Specifically the present study investigated the influence of mental flexibility and effortful control on children's responses. These

components should function together to allow children to compare multiple solutions and select the most appropriate, and to allow the child to focus their attention on key components of the task. I predict that overall children with higher scores on the executive attention tasks will be more accurate when asked about intentions. This prediction is based on our hypothesis that young children's incorrect responses may sometimes be a product of monitoring difficulties. These findings should be most robust in conditions that require children to hold conflicting representations (ie: Non Matching yes/no questions). Therefore, I predict that the format of the intentional question will influence children's accuracy. If monitoring difficulties and suggestibility lead to incorrect answers in response to questions about intentions, the forced-choice format should reduce incorrect answers because it is both less leading and directs the child's attention to the portion of the question regarding the intention of the interviewer. In essence, the child cannot simply answer "yes" or "no", but must reply with "on purpose" or "by accident". Finally, if children do in fact answer incorrectly because they are answering the descriptive outcome question instead of the intentional question, we would expect that they would falsely assent to all leading questions about events that did happen regardless of the actual intention portrayed.

CHAPTER 2

METHODS

Participants. Seventy-nine preschool children between 3 and 5 years of age (mean = 4 years, 3 months; range 3, 3 – 5, 7) were recruited for this study. The final sample was approximately 50% female (N = 40). Children's demographic information was gathered through teacher reports, and there is consequently a considerable amount of missing data concerning child ethnicity and parent education. Of the children for whom ethnicity was provided (N = 51), the majority were Caucasian (N = 41). The remainder were Asian (N = 3), Hispanic (N = 3), African American (N = 2), of mixed decent (N = 1), or "other" (N = 1). In the portion of the sample for which parental education information was provided (N = 43 for mother's education, N = 35 for father's education), nearly every parent achieved at least a college degree (N = 74) and of those, many also had an advanced degree (N = 38).

Preschoolers were recruited from preschools in and around a Northeastern research university. Permission from the preschool administrators, teachers, and children's parents were obtained before beginning data collection. As compensation for participation teachers received ten dollars at the completion of the questionnaires, and each participating classroom received a small gift of books or toys.

Apparatus. Data collection took place in quiet, isolated rooms at the preschool locations. Each room contained a table, a chair for the child, and a chair for the experimenter. Video clips were presented on laptop computers (~14" screen) and children's interviews were recorded on digital audio recorders.

Stimuli. Children were shown one of eight two-minute video clips each featuring the same adult, an eight-year-old child, and a ten-year-old child . Each of the eight videos was composed of three events which appeared to be performed either

intentionally or accidentally depending on the condition in addition to several neutral events (see Appendix A). Whether each event was performed on purpose or by accident was determined randomly and resulted in each child viewing exclusively intentional events, exclusively unintentional events, or a combination of intentional and unintentional events. The video also included other scenes of the adult and the child playing at the playground which were the same for all videos.

Materials. Executive attention was measured by the Inhibitory Control dimension of the Child Behavior Questionnaire (CBQ; Rothbart, Ahadi, Hershey, & Fisher, 2001) as well as the Dimensional Card Sort task (DCCS; Frye et al, 1995).

The CBQ Inhibitory Control dimension includes the CBQ dimensions of *Low Intensity Pleasure, Smiling/Laughter, Inhibitory Control, Perceptual Sensitivity*, and *Attentional Focusing*. For example, "When picking up toys or other jobs, usually keeps at the task until it's done", is an example of an item tapping *Attentional Focusing*. The original CBQ was slightly modified for use with children's teachers rather than caregivers. For example, questions outside of reasonable expectations of a teacher's knowledge such as "Enjoys taking warm baths" were removed. These changes resulted in a 44-item teacher-rated questionnaire, in which Items were rated on a 7-point Likert scale from "extremely untrue of this child" to "extremely true of this child". Children were given overall inhibitory control scores by reverse coding items which were negatively associated with inhibitory control and summing these reverse coded scores together with children's ratings on the rest of the items. These totals were then divided by the total number of rated items for each child's inhibitory control score. The CBQ is useful as a more global, teacher-reported measure of children's effortful control and will be compared with children's scores from the DCCS. This measure has been demonstrated to be consistent, stable, and valid-

measured by multiple converging measures of validity (Rothbart, Ahadi, Hershey, & Fisher, 2001).

The DCCS is a task measuring the executive skill of mental flexibility. In the DCCS children are presented with cards, one at a time, and are asked to sort them into two bins based on some characteristic of the card. For example, a picture of a blue triangle could either be classified into a bin with other triangles, or with other blue things. The experimenter made the sorting rules explicit to the children and changed the rules throughout the task. This task has been used to measure executive abilities, especially mental flexibility (ex: Muller, Dick, Gela, Overton, & Zelazo, 2006). Again, our hypothesis was that children with good executive skills of the sort needed to shift between relevant characteristics will be more likely to process both parts of the key questions simultaneously. Children's scores were determined by taking the sum of total correct responses and dividing by the total possible correct responses.

Procedure. There were six members of the research team, the primary investigator and five trained undergraduates. Each interviewer made an initial visit to the preschool at which they later interviewed in order to interact with the children and become familiar to them. These sessions lasted approximately one hour each.

During the first interview the experimenter chatted with the child for a few minutes to establish rapport. Once the child was seated and comfortable the interviewer introduced one of the eight brief video clips as "a movie about a grown up and a little girl playing on a playground". Immediately after viewing the video, the children were asked a randomized series of 25 questions (see Appendix B)

Children were asked questions about events that happened to confirm that they remembered seeing the event occur (3), they were asked questions about the intentions of witnessed events (3), questions about the intentions of non-witnessed events (3), yes/no questions about witnessed events with no intention (6), forced choice questions

about witnessed events with no intention (6), and open-ended questions about the intentional events that did happen (3). All but the last question category included exclusively yes/no, 2-option forced choice, or a combination of the two question types.

The subset of questions which asked about the intentions of witnessed events is of primary interest in the present study. Within this category each child was asked three question types. The first type is a "Match" question. Match questions were yes/no questions in which the correct answer should be "yes". In other words, if a child was asked "Did the grown up bump into the little girl on purpose?" and the grown up did in fact bump into the little girl on purpose, the correct answer would be "yes" and this would qualify as a "Match" question. The prompt in the question matches the event. NonMatch questions were just the opposite. These were yes/no questions where the answer should be "no". The prompt did not match the event. The final question type about witnessed, intentional events was "Forced Choice". These were questions in which children were presented with two options and asked to choose between them. For questions about intentions these options were "on purpose" or "by accident". For example, "Did the grown up bump into the little girl on purpose or by accident".

These last three questions in the interview were open-ended and were intended to give children the chance to elaborate on their responses and explain their answers. Though very few children added new information, and of those that did, the new information was often incorrect or irrelevant.

In the second visit, interviewers administered the DCCS. We also asked the children's teachers to complete a demographic form and the CBQ for each participating child.

CHAPTER 3

RESULTS

Analysis Plan

Initially, a Generalized Estimating Equation (GEE) was performed with children's accuracy as the response variable. This model included Intentionality of Question (Intention vs. Non Intention question), Question type (Match, NonMatch, Forced Choice), and children's age (~3-4 years, ~5-6 years) as independent variables along with all combinations of interaction effects with Intentionality of Question. The finding of this analysis was a main effect of Intentionality of Question ($x^2 = 31.15$, $p < 0.01$). Children were significantly less accurate when they were asked questions about intentions. Figure 1 depicts this difference by presenting the mean responses for each intentionality category. For the remainder of the analyses intentions questions and non intentions questions were examined separately in order to more sensitively determine patterns within intention and non intention questions.

Accuracy and Question Type

Non Intention Questions. Data were subjected to a GEE with children's accuracy as the response variable. The original model included question type (Match, NonMatch, and Forced Choice), question order (four randomly designed question orders), video type (8 versions), CBQ score, DCCS score, age, and gender. Preliminary analyses demonstrated that there was no main effect of age, gender, question order, video type, and no effects of CBQ score or DCCS score. The final model was reduced to include only question type, question type*age, CBQ score and DCCS score. The executive attention measures were kept in the model because though they were not significant, they are of important theoretical significance. Specific comparisons were done through the differences of least square means. The right side of figure 1 shows the differences

13

between the means by question type. Figure 2 illustrates the interaction between question types.

The main findings are that children appear to be equally accurate at matching yes/no questions and 2-option forced choice questions, they are significantly better at matching yes/no questions as compared to non matching yes/no questions ($x^2 = 7.43$, $p < 0.05$) and significantly better at forced choice questions as compared to non matching yes/no questions ($x^2 = 8.37$, $p < 0.05$). Linear contrasts revealed that children are no better at forced choice questions than they are at yes/no questions overall (ie: matching and non matching yes/no questions combined).

However, the pattern is marginally different for older children (range = 4yrs, 5mo; 5yrs, 7mo; N = 32) than for younger children (range = 3yrs, 11mo; 4yrs, 3mo; N = 46) ($x^2 = 5.35$, $p = 0.07$). There was no effect of question type for older children. In other words, for older children, accuracy for Match questions was no better than for any other question types, and their accuracy for Forced Choice questions was no better than for Non Match questions. Therefore, the younger children appear to be driving the overall effects of question type. The younger children were more accurate at Matching questions than Non Matching questions ($x^2 = 7.75$, $p < 0.05$), and more accurate at Forced Choice questions than they were at Non Match questions ($x^2 = 11.32$, $p < 0.01$). For example, a child would be more accurate when asked a question such as "Did the little girl sit on the swing?" because the correct answer was "yes" (the little girl actually did sit on the swing), than if the child was asked a question such as "Did the grown up have a red jacket?" (the jacket was actually black). A child would also be more accurate when asked "Did she go on the swings first, or down the slide first?" compared to when the child was asked a yes/no question to which the answer should be "no". However, older children are not significantly better at any one

question type than are younger children (see figure 2). This may be because younger children have increased accuracy in the Match condition because of the "yes" bias.

Intentions Questions. In order to look at intentions questions separately, these data were also subjected to a GEE with children's accuracy as the dependent variable. The original model for intentions questions included question type (Match, NonMatch, and Forced Choice), intention of the original event (on purpose, by accident), question order (four randomly designed question orders), video type (8 versions), CBQ score, DCCS score, age, and gender. Preliminary analyses demonstrated that there were no effects of age, gender, question order, video type, or executive attention measures. It should be noted that the model for the intentions questions did not show the age x question type interaction that the non intention model demonstrated. The final model was reduced to include only question type, intention of the original event, CBQ score and DCCS score. Specific comparisons were done through the differences of least square means. The left side of figure 1 shows the differences between the means by question type. Figure 3 illustrates that there was no interaction between age and question type for intentions questions.

The overall pattern of results for question type was similar to that of non intentions questions. Children were equally accurate at Match and Forced Choice questions, and more accurate at Match questions than at NonMatch questions. While it appeared that children were also more accurate at Forced Choice questions than at NonMatch questions, the p-value became marginal after controlling for multiple comparisons. See table 1 for detailed results. Again, linear contrasts revealed that children are no better at forced choice questions than they are at yes/no questions overall.

There was also a main effect of intention in predicting child accuracy ($x^2 = 12.42$, p < 0.01). Children were more accurate overall when they were asked about events that

were performed unintentionally ($\mu = 0.53$, $\sigma = 0.50$) than when they were asked about events that were performed intentionally ($\mu = 0.39$, $\sigma = 0.49$).

Response Bias

Discrimination and Criterion estimates were calculated for each child's performance on yes/no and forced choice questions without intention, as well as for children's responses overall (across children) for each of the intentions questions (yes/no and forced choice) using Signal Detection Theory (Green & Swets, 1966). I was limited in the types of discrimination and criterion calculations I could conduct because each child got only one matching intention question, one non matching intention question, and one forced choice intention question. This design was limited by the presumed limited attention spans of preschool aged children precluding the use of a lengthier testing session. Therefore, criterion scores were used to approximate children's biases, while their raw scores were used for the bulk of statistical analysis.
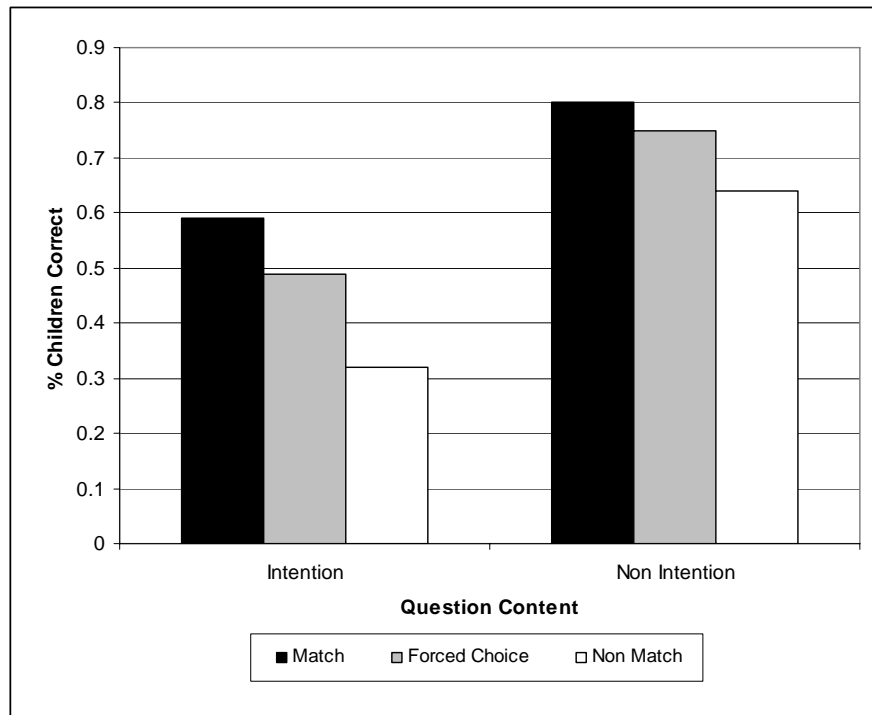

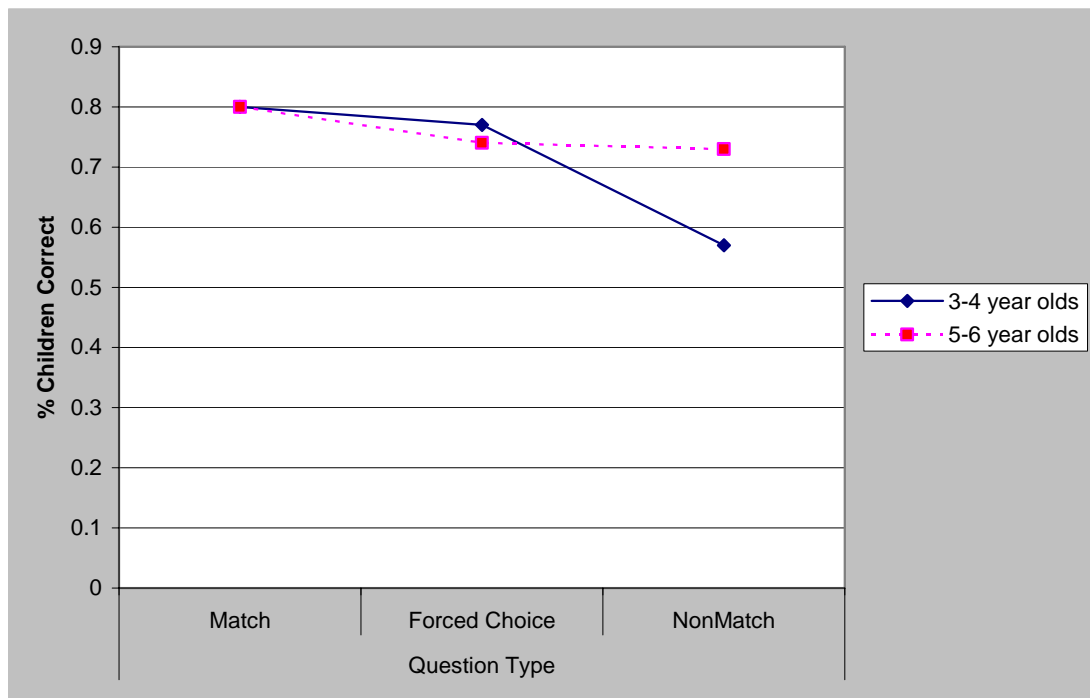**Figure 1. Percent of children correct as a function of question content and question type.**

Figure 2. Percent of children correct as a function of the interaction of question type and age for non intentions questions.
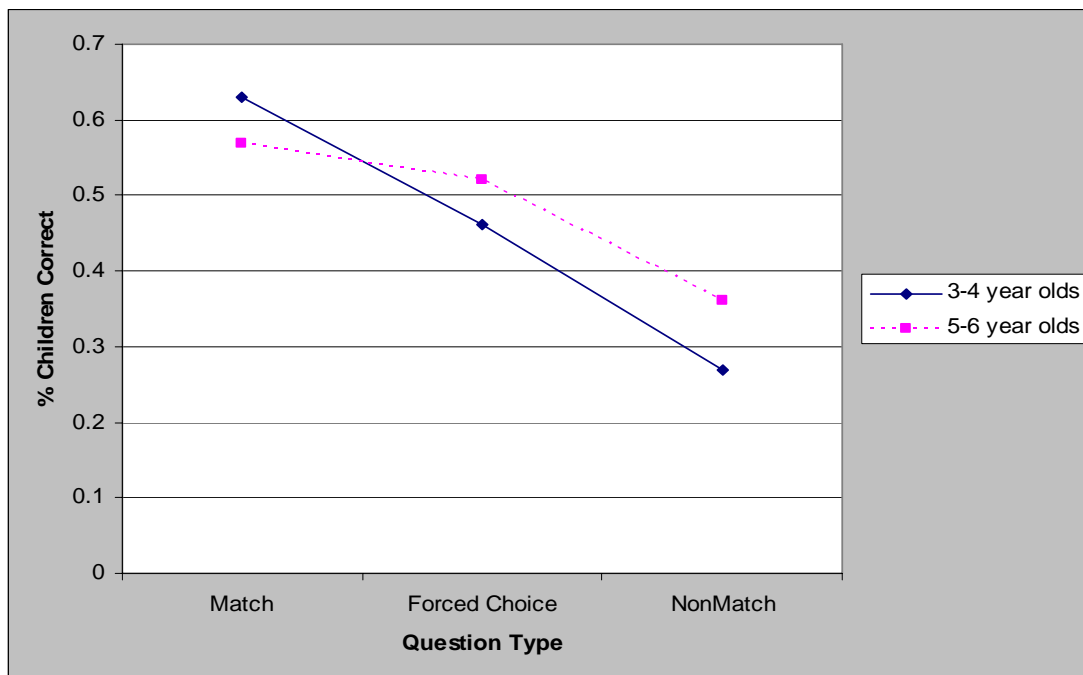
Figure 3. Percent of children correct as a function of the interaction of question type and age for intentions questions.

Table 1. Results of least square means predicting accuracy from event actual intention and question type for intentions questions.

Variable	Intention	Question Type	Intention	Question Type	Estimate	χ^2	P-value
Intention	Accident		Purpose		-1.07	12.42	<0.01
Question Type		Forced Choice		Y/N-yes	0.48	2.01	0.16
Question Type		Forced Choice		Y/N-no	-0.79	4.55	0.03
Question Type		Y/N-yes		Y/N-no	-1.26	10.15	<0.01

Overall, children were biased to respond "yes" to yes/no questions, and to respond with the second option when asked forced choice questions. The "yes" bias was consistent for both non intention questions (d' = 2.88, C = -0.52) and intentions questions (d' = -0.54, C = -0.50), as was the second position bias (intention: d'fc = 3.56, C = 0.52; non intention: d'fc = 0.91, Cfc = 0.46). For example, if a child was asked, "Was the climbing structure made of plastic?" she would be biased to respond "yes". If a child was asked "Did the little girl go on the swings first, or did she go down the slide first", the child would be biased to answer that the little girl went down the slide first. Though the assent bias cannot be regressed with age for the intentions questions because it is approximated by a single score, there was a significant interaction between yes bias and age for the non intentions questions. The "yes" bias

decreased with age, and in fact, the oldest children in the sample exhibited a moderate "no" bias (β= .22, p = 0.05). Children were also biased to choose the second option (by accident) when they were asked about the intentions in an event which they never saw (t = 4.66, p < 0.01). For example, a child would be more likely to respond that the adult stepped on the flower by accident, than they would be to respond that the adult stepped on it on purpose, when the child did not see the adult step on a flower. A logistic regression with accuracy as the dependent variable and age as the independent, continuous variable demonstrated that children's responses to this question did not differ by age.

The Non Match Condition

Analyses were also conducted to determine whether the children who performed well in the Non Match condition were different from the children who performed poorly in this condition. While performance on Non Matching non intention questions did not predict children's performance on intention questions, CBQ did predict their performance on Non Matching non intention questions ($\chi^2 =$ 7.39, p = 0.01).

CHAPTER 4

DISCUSSION

Accuracy and Question Type

This study provides further evidence that the form and content of the questions that interviewers pose to young children influences children's response accuracy. The findings that accuracy levels differed between non intentions and intentions, and that accuracy levels varied by question type in both groups illustrates this point.

That children have difficulty reliably answering questions about intentions is in line with previous research, however, children in this study were not just unreliable, they were significantly below chance. In fact, even in the Match condition, when response bias should lead to higher accuracy, children were performing no better than chance. This is especially surprising considering that there were no age effects for the intentions questions and that the oldest children in this sample were nearly six years old.

The dramatic decrease in accuracy between the non intentions questions and the intentions questions is likely a product of the language of intention. While previous research is fairly convincing that very young children and infants have some concept and perception of intentionality, it appears that children have extreme difficulties accurately using the language of intention even late in the preschool years. This is especially surprising considering earlier observational evidence that preschool-aged children use intentional language, and evidence that they respond credibly to intentions questions (Kulkofsky et al, 2007).

While there were significant effects of question type in this study, these effects wash out when yes/no questions are combined and contrasted with 2-option forced choice questions. In other words, in forensic settings, where the interviewer would not

21

in fact be aware of whether or not the correct answer to a yes/no question should actually be "yes" or "no", there is no accuracy benefit to asking young children yes/no questions as compared to forced choice questions, or vice versa.

It is also of note that for the non intentions questions only the younger children were influenced by the type of question posed to them. Older children were equally accurate at all three types. The obvious explanation would seem to be that the older children would have accuracy scores closer to ceiling, and thus we would not be able to detect variability between the question types. However, this was not the case. The older children were no more accurate at any type of non intention question than the younger children, which is interesting in itself. This is likely an effect of differing bias scores. Whereas older children's accuracy was consistent across question type, young children's "yes" biases led to lower relative accuracy in the NonMatch condition.

Response Bias

The scope of the signal detection analysis was necessarily limited in this project. Signal detection is generally used with adults due to the large number of repeated question types necessary for full analysis. Therefore, in this study discrimination and criterion estimates were calculated as approximations of response biases for this sample of children. Overall analyses were not controlled for bias. Rather, I ran the analyses with children's raw scores, and separately calculated mean bias scores across children to give a sense of children's behavior when asked about varied content using varied question types.

As presented above, in both the non intentions questions and the intentions questions, children were biased to respond "yes" to yes/no questions, and biased to respond with the second option when asked 2-option forced choice questions. The finding that children have "yes" biases when asked about intentional events is not surprising given the literature on leading and suggestive questioning (Ceci and Bruck

1993). For example, Ceci and Bruck (1993) point out that children often respond in accordance with what they perceive as the interviewer's bias. Therefore, if the question format indicates that the interviewer believes that an event happened on purpose, the child being interviewed may be more inclined to respond in accordance with this belief. However, the finding is contrary to research demonstrating that 3 year old children show no clear bias, and 4-5 year old children display "no" biases when answering questions with words they do not understand (Fritzley and Lee 2003). This discrepancy may be an effect of different methodologies, for example, Fritzley and Lee used nonsense words rather than actual English words beyond the child's abilities. Therefore, older children's "no" biases may be an understandable reaction to being asked about something they immediately recognize as a non-occurring event, namely a nonsense concept which does not exist. Thus, children's biases may change depending on whether they recognize a confusing word as a word or not.

      That children were more accurate in both conditions when answering matching yes/non questions as opposed to non matching yes/no questions is likely a product of response bias. However, response bias does not necessarily explain the finding that children were equally accurate at matching yes/no questions and forced choice questions. It may be the case that children have equal "yes" and "second position" biases, however, it may also be the case that having the correct response embedded in the question cues an increase in accurate responses. The question "Did she do it on purpose?" and "Did she do it on purpose or by accident?" both include the phrase "Did she do it on purpose". If this were true to the event, these questions formats might help children correctly respond that the event was done on purpose. However, asking the child whether the event was performed by accident provides no cue.

Accident Bias

The second position bias for the intentions question is confounded with a possible "accident" bias. Because the order of the two options in the forced choice questions were not counterbalanced we cannot know from this data whether children were biased to respond that an event they witnessed an adult perform was accidental, or whether they simply have a recency bias which is consistent with the non intentions condition and with previous research (Walker, Lunning, & Eilts, 1996, cited in Peterson & Grant, 2001). Additionally, children were significantly more likely to report that events that they did not witness the adult perform were done by accident as opposed to on purpose when given a forced choice between the two. This may be because children assume that if they did not notice, or recall the event taking place, it was likely to be unintentional. Or, it may be the case that while children do make the error of implicitly agreeing that a non-witnessed event took place, they attempt to minimize the error by claiming that the act was accidental. A final and unpredicted form of potential accident bias was demonstrated across all question types. Children were more likely to be accurate if they were asked about an event that was performed by accident. However, this effect may be partially driven by the fact that children's second position biases necessarily result in increased accuracy for forced choice questions in which the correct answer should be "accident". This is not the whole story though, because children are also more accurate at yes/no questions when they involve accidents rather than intentional events. This finding cannot be explained through biases because if children were operating solely on a yes bias they would be equally likely to answer "yes" to questions asking about accidents and questions asking about intentional events. The accident bias may have come about because children construed the events as having primarily negative outcomes, and they believed that adults do not intentionally do bad things. For example, Battin (2004) found that when children

viewed videos of an adult performing a negatively valenced event, they were more likely to falsely pluralize the number of actors throughout an interview about the event. However, when the event was positively valenced young children were more likely to give a correct, singular attribution of the actor. In other words, if the adult performed an event that was "bad" 3-5 year old children were more likely to consistently say that "they" did it without further specification, when in fact only one person performed the act. This may be an indication that children attempt to diffuse the blame behind negatively perceived acts performed by adults. Or, it may be just the opposite. Leslie, Knobe, and Cohen (2006) found that preschool aged children tend to judge negative outcomes as being products of purposeful acts, and positive outcomes to be products of accidental acts. However, in the Leslie et al study the actors were toy props, and not an adult as in the present study. Further study is required to elucidate this finding.

Executive Attention

Contrary to my original hypothesis, neither measure of executive attention had a reliable effect on children's accuracy in either content condition. Executive attention was predicted to influence children's accuracy through two separate mechanisms. The first process was through focusing of attention during the original task, the second was through superior question comprehension and conversational abilities. However, it seems that either these convergent measures of executive attention are tapping something other than monitoring and attention, or that children do not rely on these abilities in order to retain information and respond to questions in the short term. A possible explanation for the overall null findings for the non intentions questions may simply be that for the most part, children performed quite well at these questions. Perhaps there was not enough variability in overall accuracy for executive attention to have much impact, bearing in mind that these children all came from middle class

families and were probably all average or above average intellectually. Perhaps these measures failed to predict because of truncated range. This will require future research with broader ranges of family SES. However, executive attention was even less predictive of children's accuracy on intentions questions. This implies that children's difficulties in answering intention questions goes beyond monitoring. The explanation may be as simple as vocabulary. Perhaps preschool-aged children are not yet able to reliably understand the meaning behind intentional language. However, this is worth further exploration, especially given our preliminary data demonstrating that children spontaneously use intention language (Kulkofsky et al, 2007).

Non Match Condition

The Non Matching condition was singled out because it is theoretically the most challenging condition for preschool aged children. Non Match questions require a child to maintain conflicting representations, one representation of the actual event, and one of the alternatives posed in the interview question. That children's accuracy on the Non Matching non intentions questions did not predict their scores in the intentions conditions is somewhat surprising because one would expect that a child who performs well at the most complex question type would also perform comparatively well overall. However, it seems to be the case that in the intentions conditions particularly, children are driven by response biases rather than by a systematic answering strategy. Given this possibility, it makes sense that their scores in a more cognitively controlled category would not necessarily map onto their guessing responses. Additionally, if they are using an unidentified strategy beyond bias, it is ineffective and unlikely to map onto the effective strategies they are using to answer non intentions questions.

While responses on Non Match non intention questions did not predict overall accuracy in the intentions category, children's Non Match non intention accuracy was

the only factor that was predicted by an executive attention measure. Again, if children were guessing randomly in the intention conditions it makes sense that executive attention measures would not predict their performance there. It also follows logically that because Non Matching questions require children to monitor conflicting representations, executive attention would be more predictive of this question type than any other in the non intentions condition.

Implications

It seems that whether children are asked questions about intentionality or not, they are biased to respond "yes" to yes/no questions, and biased to respond with the second option when asked 2-option forced choice questions. These biases emphasize the impact that interviewers have on children's accuracy. For the most part, children's accuracy hinges on the types of questions they are asked. However, in forensic settings any benefits for the match question type are washed out by the fact that actual investigators do not know the accuracy of the question they are asking and so in real life Match and Non Match yes/no questions should be considered equivalent. In this study, children were equally accurate when they were asked forced choice questions as compared to when they were asked yes/no questions overall. This is a positive finding for the non-intentions questions because children were quite accurate overall when yes/no questions were combined.

However, though the effects of question type were even more pronounced in children's responses to questions about intentions, children's accuracy was only at chance in the most accurate condition. Importantly, this accuracy level was achieved when children were asked matching yes/no questions, which interviewers cannot control for in the field. Further, the propensity for children to respond "yes" to yes/no questions led to a staggering number of false assents. False assents can be extremely dangerous in interviews. For example, a false assent may imply guilt in a case of

accidental touching or injury. Further, assents in general shape the way interviewers perceive a case and influence the direction of further questioning. This may happen because a child's assent to a question implying intention may give the interviewer an "intention" bias, which has been demonstrated to lead to biased interviews (Bruck & Ceci, 1999). For example, if a child falsely assented that, for example, a caregiver intentionally injured them, the interviewer would be likely to incorporate this false information into further questioning, possibly cementing it into the child's testimony.

Therefore, in cases in which intent may be at issue, it may be important to establish a specific child's competency at answering questions about intentions. Currently, no state requires or recommends this precaution. Further, this study confirms previous research that verbal measures are likely to provide low intentionality competency estimates for preschool-aged children. A direct application of the results of this study may be to encourage those charged with interviewing children in forensic settings to question children about intentions nonverbally. For example, Lyon has done work on another forensically relevant abstract concept, children's understanding of truth and lies (Lyon, 2000, 2002; Lyon & Saywitz, 1999, 2000). In this work he has demonstrated that children are far more accurate at demonstrating their competency in this area when they were given visual scenarios and provide behavioral responses to those scenarios. This is also in line with the infant research mentioned earlier, in which children were able to demonstrate some intentional understanding through gesture and eye gaze (ex: Legerstee & Barna, 2000; Meltzoff, 1995).

In summary, researchers and forensic interviewers alike must be aware that while children may sound competent to answer abstract questions, such as those concerning intentions, their accuracy performance on these questions can be quite poor, at times reliably below chance. Though children can be highly accurate when

answering yes/no and forced choice questions, they are also influenced by the form and content of the questions asked. We must be aware of these vulnerabilities when interviewing children in any capacity because, particularly in the context of intentional understanding, there is a risk of greatly underestimating children's ability to understand intentionality, or, in converse, overestimating their competency to verbally communicate about intentions.

Intention Event #1: The adult either intentionally or unintentionally bumped into the child. In the intentional condition the adult looked directly at the child as she walks towards her and once she has bumped into the child the adult said "Ha! There!" In the unintentional condition the adult appeared distracted and did not look at the child until after she bumped into her. At this point, the adult will said "Oops! Are you OK?".

Intention Event #2: The adult either intentionally or unintentionally kicked a soccer ball. In the intentional condition the adult ran about 15 feet, then kicked the ball towards the child and another girl. At this point she said "Wow! Look how far that went!". In the unintentional condition the adult again ran about 15 feet, and tripped on the ball, sending it towards the two girls. This time she looked surprised and said "Oops!" and the two girls jumped out of the way of the ball.

Intention Event #3: The adult either intentionally or unintentionally spilled a cup of water. In both video segments the adult filled a cup with water from a water bottle. She then placed the cup in front of her on the table, looked away, and drummed her fingers on the table. In the intentional condition she then looked at the cup and knocked it off the table with her hand. Then, she laughed and pointed at the spilled water. In the unintentional condition the adult started a reaching motion and knocked the cup off of the table in the process. At this point, she covered her mouth with her hand, said "Oh no!" and looked at the spilled water in surprise.

APPENDIX B

Events that did happen:

    1) Did the grown up bump into the little girl?

    2) Did the grown up kick the soccer ball?

    3) Did the grown up spill the water?

Events that did happen with matching intention, mismatching intention, or forced choice:

    4) Did the grown up bump into the little girl on purpose/by accident/on purpose or by accident?

    5) Did the grown up kick the soccer on purpose/by accident/on purpose or by accident?

    6) Did the grown up spill the water on purpose/by accident/on purpose or by accident?

Intentions questions about things that did not happen:

    7) Did the grown up throw the basketball on purpose?

    8) Did the grown up drop the doll by accident?

    9) Did the grown up step on the flower on purpose or by accident?

Yes/No response bias controls (first three, answer: yes, second three, answer: no):

    10) Did the little girl play on the swings before the slide?

    11) Did the little girl sit on the swing?

    12) Did one of the girls slide down the handrails of the stairs?

    13) Was the climbing structure made of plastic?

    14) Did the grown up have a red jacket?

    15) Did the grown up fall off the climbing structure?

Forced choice response bias controls:

    16) Did the two kids come with the grown up or meet her there?

17) Who was taller, the grown up or the little girl?

18) Did the grown up get to the park first, or did the little girl?

19) Did the little girl have brown or blond hair?

20) Did the grown up push the little girl on the swing, or did the little girl push the grown up?

21) Did the grown up go down the slide or did the little girl go down the slide?

22) Was the grown up nice or mean?

Open ended questions

23) What else can you tell me about when the grown up bumped into the little girl?

24) What else can you tell me about when the grown up kicked the soccer ball?

25) What else can you tell me about when the grown up spilled the water?

REFERENCES

Astington, J.W. (1999). The language of intention: Three ways of doing it. In P.D. Zelazo,
J.W. Astington, & D.R. Olson (Eds.), *Developing theories of intention: Social
understanding and self-control* (pp. 295-315). Mahwah, NJ: Erbaum.

Baird, J.A. and Moses, L.J. (2001). Do Preschoolers Appreciate That Identical Actions
May Be Motivated by Different Intentions? *Journal of Cognition and
Development*, 2(4), 413-448.

Battin, D.B. (2004). *What does it mean to understand and answer simple questions?
Young children's evidentiary testimony within the paradigm of discourse
pragmatics*, Doctoral dissertation, Cornell University, August 2004.

Battin, D.B. and Ceci, S.J. (2003). Children as Witnesses: What We Hear Them Say
May Not Be What They Mean. *Court Review*, 4-5.

Behne, T., Carpenter, M., Call, J., Tomasello, M. (2005). Unwilling Versus Unable:
Infants' Understanding of Intentional Action. *Developmental Psychology*,
41(2), 28-337.

Brady, M.S., Poole, D.A., Warren, A.R., and Jones, H.R. (1999). Young Children's
Responses to Yes-No Questions: Patterns and Problems. *Applied Development
Science,* 3(1), 47-57.

Bruck, M., and Ceci, S.J. (1999). The Suggestibility of Children's Memory, *Annual
Review of Psychology*, 50, 419.

Carter, C.A., Bottoms, B.L., & Levine, M. (1996). Linguistic and socioemotional
influences on the accuracy of children's reports. *Law and Human Behavior,*
20(3), 335-358.

Ceci, S. J. and Bruck, M. (1993). "Suggestibility of the Child Witness - a Historical
Review and Synthesis." <u>Psychological Bulletin</u> 113(3): 403-439.

Ceci, S.J., Ross, D.F., Toglia, M.P. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General*, 116(1), 38-49.

Choi, S. (1991). "Childrens Answers to Yes No Questions - a Developmental-Study in English, French, and Korean." <u>Developmental Psychology</u> 27(3): 407-420.

Feinfield, K.A., Lee, P.P., Flavell, E.R., Green, F.L., and Flavell, J.H. (1999). Young Childrens' Understanding of Intention. *Cognitive Development*, 14, 463-486.

Fritzley, V. H. and K. Lee (2003). "Do young children always say yes to yes-no questions? A metadevelopmental study of the affirmation bias." *Child Development* 74(5): 1297-1313

Frye D, Zelazo PD, Palfai T. 1995. Theory of mind and rule-based reasoning. *Cognitive Development,* 10: 483-527.

Green, D. M. and Swets, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.

Hamlin, J.K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450(22), 557-559.

Kulkofsky, Klemfuss, & Ceci (2006). [The relationship of creativity and sociability with suggestibility in repeated interviews]. Unpublished raw data.

Legerstee, M., J. Barna, et al. (2000). "Precursors to the development of intention at 6 months: Understanding people and their actions." <u>Developmental Psychology</u> 36(5): 627-634

Leslie, A.M., Knobe, J., and Cohen, A. (2006). Acting Intentionally and the Side-Effect Effect. *Psychological Science*, 17(5), 421.

Lyon, T. D. (2000). "Child witnesses and the oath: Empirical evidence." <u>Southern California Law Review</u> 73(5): 1017-1074.

Lyon, T.D. (2002). Child Witnesses and the Oath. In H.L. Westcott, G.M. Davies, &

R.H.C. Bull (Eds.), Children's testimony: A handbook of psychological research and forensic practice (pp.245-260). West Sussex, UK: John Wiley.

Lyon, T.D., & Saywitz, K.J. (1999). Young Maltreated Children's Competence to Take the Oath. *Applied Developmental Science*, 3(1), 16-27.

Lyon, T.D., & Saywitz, K.J. (2000). Qualifying Children to Take the Oath: Materials for Interviewing Professionals.

Meltzoff, A.N. (1995). Understanding the Intentions of Others: Re-Enactment of Intended Acts by 18-Month-Old Children. *Developmental Psychology*, 31(5), 838-850.

Moses, L. J. (1993). Young children's understanding of belief constraints on intention. *Cognitive Development, 8*, 1–25.

Moses, L.J., Baldwin, D.A., Rosicky, J. G., Tidball, G. (2001). Evidence for Referential Understanding in the Emotions Domain at Twelve and Eighteen Months. *Child Development*, 72(3), 718-735.

Muller, U., Dick, A.S., Gela, K., Overton, W.F., & Zelazo, P.D. (2006). The role of negative priming on preschoolers' flexible rule use on the Dimensional Change Card Sort Task. Child Development, 77(2), 395-412.

Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development,* 67, 3045-70.

Peterson, C., Dowden, C., & Tobin, J. (1999). Interviewing preschoolers: Comparisons of yes/no and *wh*- questions. *Journal of Law & Human Behavior*, 23, 539-556.

Peterson, C. and M. Grant (2001). "Forced-choice: Are forensic interviewers asking the right questions?" <u>Canadian Journal of Behavioural Science-Revue Canadienne Des Sciences Du Comportement</u> **33**(2): 118-127.

Piaget, J. (1932). *The moral judgement of the child*. Harcourt, Brace.

Poole, D.A., & Lindsay, D.S (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*, 60, 129-154.

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at 3-7 years: The Children's Behavior Questionnaire. Child Development, 72, 1394-1408.

Schult, C.A. (2002). Children's understanding of the distinction between intentions and desires. *Child Development*, 73(6), 1727-1747.

Shultz, T.R., Wells, D., & Sarda, M. (1980). Development of the ability to distinguish intended actions from mistakes, reflexes, and passive movements. *British Journal of Social and Clinical Psychology*, 19, 301-310.

Shimizu, Y.A., Johnson, S.C. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science*. 7(4), 425-430.

Smith, M.C. (1978). Cognizing the Behavior Stream: The Recognition of Intentional Action. *Child Development*, 49, 736-743.

Steffensen, M.S. (1977). Satisfying inquisitive adults: Some simple methods of answering yes/no questions. *Journal of Child Language*, 5, 221-236.

Walker, N.E., Lunning, S.M., & Eilts, J.L. (1996). Do children respond accurately to forced-choice questions? Yes or no. Presented at the NATO Advanced Study Institute on Recollections of Trauma: Scientific Research and Clinical Practice. Talmont Saint Hilaire, France.