Spring 5-17-2022

Children's ability to understand and respond to wh- questions about the mechanics of abuse

Kaileigh P. Conti
CUNY John Jay College, kpconti5@gmail.com

Children's ability to understand and respond to wh- questions about the mechanics of abuse

A Thesis Presented in Partial Fulfillment of the Requirements for the Degree of
Master of Arts in Forensic Psychology
John Jay College of Criminal Justice
City University of New York

Kaileigh Conti

May 2022

Children's ability to understand and respond to wh- questions about the mechanics of abuse

Kaileigh Conti

This Thesis has been presented to and accepted by the Office of Graduate Studies, John Jay College of Criminal Justice in Partial Fulfillment of the Requirements for the Degree of Master of Arts in Forensic Psychology

Thesis Committee

Thesis Advisor: Kelly McWilliams, Ph.D.

Second Reader: Charles B. Stone, Ph.D.

Third Reader: Emily Haney-Caron, Ph.D.

Abstract

One type of forensically-relevant information that can be difficult to obtain is that pertaining to the "mechanics of abuse." More specifically, information that includes the descriptions of body positioning and clothing placement. Generally, the recommended strategy for questioning children in legal and forensic settings is to use broad invitations (e.g., "Tell me everything that happened") and wh- questions (e.g., who, what, where, when, why, how). However, when it comes to the mechanics of abuse, there is some conflicting evidence. Some research suggests open-ended wh- questions are best in cases where the mechanics are hard to describe (e.g., intermediate clothing placement) (Stolzenberg et al., 2017a). However, it has also been argued that some forms of wh- questions about these topics (e.g., how questions) may be too broad or abstract for young children within the context of an interview. The current study examined 5- to 10-year-old children's responses to wh- questions about both mechanics of abuse and non-mechanics of abuse information after viewing a threatening or non-threatening stimulus to determine if wh- questions (largely "what" questions) are too broad or ambiguous for young children. Results indicate that young children are able to respond to wh- questions about these topics with few expressions of confusion; however, age differences did emerge suggesting these questions may need to be used with caution when questioning very young children (i.e.., 5 to 6 years old and younger).

Keywords: Forensic interviewing, Child abuse, Mechanics of abuse

Children's ability to understand and respond to wh- questions about the mechanics of abuse

One of the most prominent issues involved with the identification, investigation, and prosecution of cases of child maltreatment is that often the strongest piece of evidence available is the testimony of a young child (Myers, 1989). Although children are capable of providing reliable reports and can be convincing witnesses, they are also more vulnerable to misinformation, suggestion, and confusion than adult witnesses (Stolzenberg et al., 2017a). As a result, the quality of forensic interviewing and trial questioning becomes paramount in these cases. If a questioner does not use developmentally friendly, empirically based techniques, they could cause a child's testimony to be deemed incredible and have the case dropped, or elicit a false narrative from the child and charge an innocent person. For example, skepticism about a five-year-old's testimony arose during State v. Emmett (1992), when a young witness became confused about his and his father's clothing placement at the time of an alleged sodomy during a pre-trial hearing. More specifically, when asked whether his clothes were "on or off" during the alleged assault, the child responded that his clothing was "on." This simple response led the adult factfinders to believe that the child's testimony was not credible evidence, as they believed it was impossible to sodomize a child with his clothing on (State v. Emmett, 1992). There was no consideration given to the possibility that the child's clothes may have been partially on or pulled down (Stolzenberg et al., 2017a). Given that the boy's testimony was the only direct piece of evidence and was no longer viewed as credible, the trial judge was forced to drop the charge due to a lack of evidence.

The above case example both emphasizes the importance of quality of questions with young children and highlights another important topic, questioning children about the

"mechanics of abuse." The mechanics of abuse (Sullivan et al., 2021) refers to information about body movement, body placement, facial expressions, and clothing placement in cases of abuse. Questioning children about the mechanics of abuse is important because it can provide information that assists in determining the difference between abusive contact and typical caregiving (e.g., bathing) or determining the exact type of abuse and specific charging (e.g., force, penetration) (Lyon et al., 2017). However, questioning children about these topics can be particularly difficult because this information often does not spontaneously arise in children's narrative responses (Stolzenberg & Lyon, 2017), and can be difficult to ask about directly because the topics may go beyond a child's scope of knowledge (e.g., penetration).

Research has repeatedly shown that open-ended questions, including wh- questions (who, what, where, when, why, how), are ideal for interviewing children (Wylie et al., 2020). Empirically based interview recommendations regularly warn against the use of closed-ended questions (i.e., yes/no, forced-choice questions). One reason is because of formal reticence, or children's tendency to provide as brief a response as possible, which leads to responses lacking detail (Stolzenberg & Lyon, 2014). Another is that there is a heightened risk, compared to open-ended questions, that these questions and answers will contain false information (McWilliams et al., 2021). However, when legal professionals need specific pieces of information, they need guidance on the exact type of wh-question to ask. Researchers have just begun to examine the best strategy for obtaining information about the mechanics of abuse. Stolzenberg and colleagues (2017a) recently examined young children's abilities to describe both simple (e.g., fully on/off) and intermediate (e.g., pulled down) clothing placement. Their findings suggest that wh-questions (e.g., "where were his clothes") are ideal for eliciting full descriptions of clothing placement from young children (3 to 6 years old), providing early evidence that wh- questions

can be used to obtain mechanics of abuse information. However, another study by Henderson and colleagues (2022) found that how questions about the mechanics of abuse created confusion in witnesses, suggesting that the general recommendation of "using wh- questions" is likely too simplistic.

NICHD Investigative Interview Protocol

Many procedures and guidelines for forensic interviews have been developed in recent years in an effort to promote the best method of practice. The most well-researched protocol, the National Institute of Child Health and Human Development (NICHD) Investigative Interview Protocol, is often considered the "gold" standard (Benia et al., 2015). This protocol was created with consideration of child interviews and suggests following a specific format that begins with an introductory phase and is followed by an allegation phase which focuses on eliciting detailed, but reliable, disclosures of wrongdoing. In the introductory phase, it is important the interviewer relays the "ground rules," which are a specific set of instructions that help the child understand how they should respond during the interview. The ground rules include telling the child that they are able to say remarks such as "I don't remember" and "I don't know," or correct the interviewer when appropriate (Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007). Following the ground rules, interviewers are instructed to engage in specific rapport-building strategies. The rapport building includes asking children to talk about neutral topics, such as what the child likes to do or what happened on the child's last birthday. The goal of this phase is to make children comfortable, but it is also used to practice the types of questions and responses that will be used during the allegation phase. The recommended course of questioning in rapport building and throughout the interview begins with free-recall prompts known as invitations, ("Tell me everything…"). These prompts are used to elicit narrative responses from children that will guide

the direction of the interview. Following invitations, interviewers are supposed to follow up using cued invitations, which are prompts that expand on the child's narrative response using only information that was provided by the child (e.g., "You said X, tell me more about X."). Once the interviewer has exhausted cued invitations, they can use wh- questions (who, what, when, where, how), to attempt to elicit additional details or clarifications. If vital details are still missing, the interviewer then is advised to continue with limited option-posing questions (yes/no or force-choice questions) (Lamb, Orbach, Hershkowitz, Horowitz, et al., 2007), but it is important that these questions are used with caution and only when necessary.

The structure of the NICHD protocol, including the recommendations about specific question types, is based on decades of research examining children's abilities as witnesses (Lyon et al., 2019). The different phases and particular patterns of questions are specifically designed to minimize the potential for suggestive influences and misinformation, attenuate children's tendencies to guess, and provide children with the best opportunity to reliably account for what they know (Lyon et al., 2019). Lamb and colleagues (2007) tested the NICHD Investigative Interview Protocol's validity in children and discovered it can be used on children as young as four years of age to elicit accurate and detailed information. Additionally, Lamb and colleagues (2007) have found that when using the NICHD protocol, interviewers elicit more detailed and accurate information as compared to when they question children without a standardized protocol (Lamb, Orbach, Hershkowitz, Esplin, et al.et al., 2007; Orbach et al., 2000)

A main tenet of structured forensic interview protocols is optimizing the use of free-recall prompts, or invitations. These prompts are open-ended questions that have the child recall information following an incident, they can be posed as a statement, question, or imperative. Open-ended questions are preferred while conducting interviews because they require recall

from memory, whereas closed-ended questions often rely on the interviewee recognizing an option provided by the interviewer (Lamb, Orbach, Hershkowitz, Esplin, et al., 2007). In terms of accuracy, open-ended questions in forensic interviews tend to be more accurate than more focused or closed-ended prompts (Lamb, Orbach, Hershkowitz, Esplin, et al., 2007; Lamb & Fauchier, 2001; Orbach & Lamb 1999, 2001). Additionally, responses to free-recall prompts are about three to five times more informative than those of closed-ended prompts (Lamb et al., 1996; Sternberg et al., 1996; Sternberg et al., 2001). Therefore, they are typically the most recommended type of question when interviewing young children. Unfortunately, an interviewer may not get all the information they need from a child witness in response to an invitation, thus interviewers often need to ask more specific questions. When this is the case, interviewers are told to turn to wh- questions (Stolzenberg & Lyon, 2017). Wh- questions are typically preferred over closed-ended questions because, when phrased correctly, these questions still pull for informative responses, do not include presupposition or suggestion, and do not proffer an acceptable answer in the structure of the questions (McWilliams et al., 2020). However, not all wh- questions are created equal; some wh- questions are more like closed-ended questions than they are open-ended questions. For instance, wh- questions about topics that are learned by rote (e.g., colors, numbers) can fall prey to many of the same pitfalls as closed-ended questions. This is because children learn by memory. The explanation for this is that when children are asked about a concept for which they have a lexical domain, they can easily draw from a list and come up with a reasonable response. Additionally, wh- questions about difficult topics can become vague, use language that is adult-centric, or even include suggestion. Thus, while the use of wh-questions is a common recommendation to elicit certain forensically-relevant information from children, how a wh-question is worded is important.

Wh- Questions and Mechanics of Abuse

Determining how to question children about the mechanics of abuse is relevant because most forensic interviews (72%) include questions about this topic, and children are questioned, on average, five times per interview about these topics (Stolzenberg & Lyon, 2017). Interestingly, only 29% of the mechanics of abuse information that arises during interviews is from spontaneous mentions (Stolzenberg & Lyon, 2017), suggesting that children are not bringing up the information themselves, but rather are providing it in response to interviewer questions. Additionally, spontaneous descriptions of clothing placement and other mechanics information occurred mainly in response to wh- questions (86%, otherwise occurring in response to yes/no questions), meaning the children are not giving the information during their narrative following an invitation. These findings suggest that collecting data on children's ability to answer wh- questions about the mechanics of abuse is imperative because interviewers are regularly asking these questions with no empirically based guidance.

To date, the one of only known laboratory research examining children's responses to questions about the mechanics of abuse is on children's ability to answer questions about clothing placement specifically. Stolzenberg and colleagues (2017a) conducted the first laboratory study examining how children respond to both closed-ended and wh- questions about clothing placement. They developed a paradigm they called the "Clothing Task," which involved figurines with a shirt and pair of shorts. During this task a research assistant manipulated the figurine's shirt and pants by opening, pulling down, or taking off the items. After each manipulation, the research assistant would ask the child a question about the clothing using either a closed-ended question ("[Child's name], are his pants on or off?") or a wh-question ("[Child's name], where are his pants?"). The manipulation of the doll's clothes at different

stages resulted in clothing placement that was either simple (completely on, completely off) or intermediate (clothing open, clothing at a low joint, clothing at midjoint). The results of this task revealed that when clothing placement was simple, the question type did not matter, children were highly accurate to both closed-ended and wh- questions. However, when the clothing placement was intermediate children were significantly more accurate in response to the wh-question. More specifically, when asked a wh-question the children would describe the clothing placement (e.g., "the shirt is on his arms"), but if they were asked the closed-ended question children would simply respond the clothes were "on" or "off." Based on these results, Stolzenberg and colleagues (2017a) concluded that wh- questions are ideal because they elicit the correct information for questions involving both simple and intermediate clothing placement.

However, anecdotally, there were some criticisms of these findings. Most commonly, when interviewers heard these findings, they would often claim that wh- questions about clothing placement specifically, and mechanics of abuse in general, are too ambiguous for children, and the children do not know what the interviewers are talking about when they ask these wh-questions. While children understood the meaning of the questions being asked in Stolzenberg and colleagues' (2017a) study, it is possible that the comprehension was due to the Clothing Task method. In their task, the only questions the researchers asked were about the figurine's clothes, and the questions were asked immediately after manipulating the clothing. Therefore, the setting and manner in which the study was conducted could have facilitated understanding of these questions. In a forensic interview, children are asked many different questions about various topics, and the interview is often conducted after a significant delay, which could result in more ambiguity.

In addition to interviewer concerns, recent research has suggested that not all types of wh- questions are easy for children to understand, especially within the context of the mechanics of abuse. Henderson et al. (2022) studied child witnesses' (4 to17 years old) difficulty with how questions during trial testimony cases of alleged child sexual abuse. The authors were interested in children's ability to understand both "how-evaluative" questions and "how-manner" questions. "How-evaluative" questions are questions that are asked about a prior or current subjective state (e.g., "How did you feel?"). "How-manner" questions are those that asked how something happened, which were often questions about the mechanics of abuse (e.g., "How did he touch you?" "How were his clothes?"). The authors predicted that children would provide more productive responses to "How-evaluative" questions than to "How-manner" questions because children learn at a young age to answer evaluative how questions (e.g., "How are you?" "How did it go?" How was your day?") but are less likely to hear and practice how questions that are asking for the manner in which an action occurred or the outcome of that action (e.g., how-manner questions). Consistent with the authors' hypothesis, the results of this study showed that "How-manner" questions produced a high rate of non-substantive responses and that children gave a significantly higher rate of non-productive responses to "How-manner" questions compared to "How-evaluative" questions.

Additionally, there was an age effect indicating that as children got older, they provided fewer non-substantive responses to these questions, suggesting the questions may become less ambiguous as children get older. More specifically, 4- to 5-year-olds gave non-substantive responses to 17% of the "How-manner" questions while 7- to 8-year-olds gave non-substantive responses to 4% of the "How-manner" questions. These age effects could reflect a variety of developmental differences that have been since in child eyewitness and memory research.

Decades of research has demonstrated that younger children are more vulnerable witnesses for a variety of reasons. They have been shown to be more suggestible, have poorer memory recall, less relevant prior knowledge that can be applied, have poorer vocabulary, and inability to comprehend complex syntax (Ceci & Bruck, 1993). Additionally, children are more prone to yield to social pressure, especially from authority figures compared to older children and adults (Ceci & Bruck, 1993). Therefore, Henderson et al.'s (2022) findings are consistent with a large body of evidence suggesting that when it comes to recounting experienced events, younger children (i.e., 6 years and younger) are in a significantly different developmental stage and, therefore, experience limitations.

Taken together the findings from Stolezenberg et al. (2017) and Henderson et al. (2022), as well as other past research, suggest that asking children about the mechanics of abuse can be difficult and that children's ability to answer these questions may develop across early to middle childhood.

The Present Study

The current study aims to add to this literature by examining children's responses to wh-questions about the mechanics of abuse. The present study expands the current literature in three important ways. First, this study expands on Stolzenberg et al. (2017a) by examining children's ability to answer questions about more topics beyond clothing placement. Second, children were asked both questions about the mechanics of abuse and questions about non-mechanics of abuse information, which is more similar to the context of a forensic interview. Third, the current research attempted to replicate the age finding observed in Henderson et al. (2022), but using a laboratory paradigm.

The present study is part of a larger study examining attention bias to threatening information and how attention bias may influence children's reports of threatening and non-threatening information. Therefore, the stimuli and questions within the study were developed for that larger purpose. We will discuss, when applicable, how this may impact our results. Our study was exploratory in nature and aimed to examine children's ability to answer wh- questions about the mechanics of abuse and non-mechanics of abuse. While we had no directional hypotheses, we did believe, based on the findings from Henderson et al. (2022), that we may see an age effect in children's ability to answer these questions.

Method

Participants

We collected data from a total of 37 children between the ages of 5-10 (M = 7.92, SD = 1.70). 60% of participants were female (n = 22) and 40% were male (n = 15), and 43% were White (n = 16), 35% (n = 13) were Asian, 11% (n = 4) were Black, 5% (n = 2) were Native American, and 5% (n = 2) were Latinx or Hispanic. Participants were recruited through posting advertisements in various Facebook groups geared towards parents or survey participation (e.g., "Westport Front Porch," "Parents of Greenwich Village," "Paid Survey Opportunities," etc.). We excluded children who did not speak fluent English, who had any diagnosed developmental delays, who were colorblind or had visual or hearing impairments, or who did not have access to a functioning computer with a camera, microphone, and stable internet connection. Once participants responded to our advertisement, we scheduled a brief screening call with them to ensure they met all inclusion criteria. In efforts to ensure children were the age their parents reported, we asked parents and children what year the child was born and noted any inconsistencies (which were rare).

Materials

Flappy Bird Game (Session 1)

"Flappy Bird" is a notoriously frustrating game that involves keeping a small bird in the air while dodging obstacles. If an obstacle is hit, the game restarts. (The game can be found here: https://flappybird.io/). We found the code for this game online and integrated it into our Qualtrics survey. The game is so difficult that it controls performance across all participants (it is not possible that any child who participates will attain a score higher than 1). Children play it for about a minute before the game stops and auto-advances to play one of the two Coach Dan speeches. After viewing the speech, they play the game again for the same amount of time. For the larger study at hand, this game served as a baseline measure memory test before children were exposed to a threat condition.

"Coach Dan" Speech (Session 1)

The main stimulus was a video of a coach giving the children a motivational speech about their "mission" (i.e., playing the game Flappy Bird). We scanned the internet for motivational speeches given by coaches in child-friendly sports movies and used these to help us draft a three-minute-long motivational speech. This study is part of a larger study investigating children's attentional bias for threat. Due to this, the "Coach Dan" speech had a threat manipulation. Although we do not have hypotheses concerning the threat condition, we believed there could be an effect of threat on children's ability to answer questions about "Coach Dan" and are therefore controlling for it in our analyses. Research has indicated that an attention bias to threatening information exists in certain populations (Bar-Haim et al., 2007), and part of that information in Coach Dan's speech included his body mechanics. If the threat was more salient,

there was a possibility children would have interpreted body mechanics differently and had different responses to our questions about the mechanics of abuse.

"Coach Dan" was played by an actor. He recorded himself giving the speech in either a "motivational" or "threatening" manner. For the "threatening" condition, we specifically instructed the actor to exhibit specific threat-related nonverbal cues. The threat-related cues were chosen from previous research and included tone of voice, posture, and facial expression (Johnson & Aaron, 2013). In the no threat condition "Coach Dan" gave the same speech and general choreography; however, his tone of voice, posture, and facial expressions were consistent with neutral cues (Johnson & Aaron, 2013). In both conditions, "Coach Dan" gave his speech in front of a large background image. The background image depicted cartoon characters having dinner.

Memory Questionnaire (Session 2)

Participants were given a memory questionnaire with two categories of questions: (1) Mechanics of Abuse questions and (2) Non-Mechanics of Abuse questions. Both categories of questions are made up of wh- questions. There were 11 Mechanics of Abuse questions, including four wh- questions about clothing placement and physical appearance questions (e.g., "What did Coach Dan do with his clothes?") and seven questions about body mechanics (e.g., "What was Coach Dan doing with his hands?"). There were seven Non-Mechanics of Abuse questions, all of which were wh- questions about the background behind Coach Dan (e.g., "What color was the room?"). See Appendix for a full list of questions.

Procedure

Initial Recruitment, Screening, and Email Communications

First, participants were recruited through advertisements, and we responded to them via email to schedule a screening call. A research assistant then called the child's parents to confirm their child was eligible to participate. If eligible, the research assistant would schedule both sessions with them, each to occur within the same week. It was important that parents and children were both available and present during both sessions. Directly after the screening call, the research assistant emailed the parent a link to a Qualtrics survey containing a detailed study description and a demographic questionnaire. We clarified that if the child was unable to read, the parent was permitted to help them fill out the form while emphasizing that the parents should try their best to not influence their child's answers or offer any input. Then, we sent two Google Calendar invitations (one for each session) and assigned them a participant code (a random number that they entered at the beginning of each session and survey they completed) which allowed us to connect their video and survey data.

Session One

The entirety of the study took place remotely over the course of two sessions. The first session took place online, unmoderated (no interaction with experimenters). Upon clicking the study link, participants were directed to our platform. When they clicked "begin survey" they were automatically prompted with a message requesting access to their camera, microphone, and screen recording. Upon completing the study, parents were given instructions on how to stop the video recording and submit the video to a secure server, from which our laboratory could retrieve it.

The session began with an animated video to explain the study and obtain consent from the parent (clicking a box to continue) and assent from the child (saying "yes" out loud) before

asking the parent to leave the room so that the child could complete the next few tasks on their own.

The child then completed a "special mission" which required them to play Flappy Bird. After a few minutes, the game temporarily stopped, and the child viewed one randomly assigned version of the "Coach Dan" Speech. They then played Flappy Bird for a few more minutes before being asked to retrieve their parent to end the session

Session Two

The second session was a video-recorded zoom conference. Research assistants were trained in a developmentally sensitive interviewing protocol (based on the recommendations from the NICHD protocol) and monitored by the study project manager. Due to the larger study, they began the session with a free-recall, narrative-style interview where they asked the child to mention everything they could remember about the Flappy Bird mission. Then, they focused in on wh- questions asking about the mechanics and non-mechanics information pertaining to Coach Dan and the background behind him.

We did not include wh- questions about Flappy Bird in our analyses because it was a different stimulus than Coach Dan's speech. We believed combining these questions could introduce a confound.

Coding

Two research assistants double-coded 20% of all responses and reached high interrater reliability on all codes (kappa > .80). All discrepancies were resolved by the PI on the project. The remaining transcripts were divided among the two researchers.

Children's responses to the wh- questions were either coded for accuracy or for "substantiveness" when the question was subjective. Accuracy codes included correct, incorrect,

or “I don’t know/I don’t remember”. Subjective codes included substantive or non-substantive. Substantive responses were those that provided logical, on-topic, and meaningful information pertaining to the question being asked. A non-substantive response would include answers that are uninformative, off-topic, simply repeating the question, or “I don’t know/I don’t remember”.

Accuracy codes were given for the Mechanics of Abuse questions about clothing placement and appearance and substantive codes were given for the questions about body mechanics. For the Non-Mechanics of Abuse questions, accuracy codes were given to all the questions except two that were about the mood of the characters in the background, those questions were given subjective codes.

There were instances in which children provided partial-incorrect answers, which were coded as such. However, we employed conservative accuracy coding and considered partial-incorrect answers as incorrect answers in our analyses. The goal of the child’s answer was less about accuracy than it was about being able to meaningfully answer the question being asked. If the child was providing some false information, we were unable to determine whether the child was not remembering correctly or not understanding the question.

For all dependent variables, codes were summed across question category and divided by the total number of questions in each question category to create proportion scores. Our descriptive analyses will provide information on the accuracy and substantive proportion scores for each question type. However, our inferential statistics are based on children’s ability to answer the question. For these variables, we summed across accuracy and substantive codes. The proportion of “answered” questions includes the sum of the accurate, inaccurate, and substantive scores divided by the total number of questions.

Results

Analysis Plan

An a priori power analysis using G*Power (Faul et al., 2007) determined that to detect an effect size of .25 with 80% power at $p < .05$ for a repeated measure ANOVA, we would need a sample size of 98 participants. Due to funding restrictions, we were only able to collect data from 37 participants.

First, we present the findings of our descriptive analyses for the mechanics of abuse and non-mechanics of abuse questions separately. We report the proportion of correct, incorrect, "I don't know/I don't remember", substantive, and non-substantive responses for each question category based on whether or not accuracy could be determined.

Following this, we examined the effects of age, threat condition, and topic of information (mechanics of abuse or non-mechanics of abuse) by conducting a repeated measures analysis of variance (ANOVA). The proportion of answered questions served as the dependent variable, age was entered as a continuous covariate, threat condition was a between subjects fixed factor, and the proportion of answered questions to mechanics of abuse information and non-mechanics of abuse information was a within factor variable.

Children's Responses to Mechanics of Abuse Questions

First, we examined children's abilities to answer the different Mechanics of Abuse questions by topic. We found that children were fairly accurate when answering clothing placement and appearance questions on average, children answered 65% of these questions correctly ($M = .65$, SD = .27). Children gave incorrect responses to an average of 20% of questions ($M = .20$, SD = .20) and tended to say "I don't know/I don't remember" to about 15% of questions ($M = .15$, SD = .19). For questions about body mechanics, we could not look at accuracy, because many were subjective. We therefore examined rates of substantive responses

and found that, on average, children gave substantive responses to 80% of these questions (*M* = .81, SD = .20) and non-substantive responses to about 20% of these questions (*M* = .18, SD = .21). When we collapsed across question type (i.e., clothing placement/appearance and body mechanics) and examined how many questions regarding mechanics of abuse children were able to answer, we see that on average children gave substantive, meaningful answers 80% of the time (*M* = .83, SD = .17) and were unable answer to answer less than 20% of the time (*M* = .16, SD = .17).

Children's Responses to Non-Mechanics of Abuse Questions

Next, we were interested in how children answered Non-Mechanics of Abuse questions. For these questions, on average, children answered slightly less than half of the questions correctly (*M* = .48, SD = .30), provided incorrect responses to 37% of questions (*M* = .37, SD = .23), and gave "I don't know/I don't remember" responses to an average of 15% of questions (*M* = .15, SD = .25). For questions about Coach Dan's mood, we could not code for accuracy as they are subjective; thus, we examined the proportion of substantive responses. We found that, on average, children were able to answer 93% of the questions about Coach Dan's mood (*M* = .93, SD = .18) and gave non-substantive responses to 5% of these questions (*M* = .05, SD = .15). When we collapsed across question types to examine children's ability to answer all Non-Mechanics of Abuse questions, we found that on average children gave substantive, meaningful answers to 91% of the questions (*M* = .91, SD = .13) and were unable to answer 8% of the questions (*M* = .08, SD = .13).

Mechanics of Abuse Questions vs. Non-Mechanics Questions

Finally, we examined whether there were any differences in children's patterns of responding across Mechanics of Abuse questions and Non-Mechanics of Abuse questions. We

conducted a repeated-measures ANOVA with the proportion of answered questions to Mechanics questions and the proportion of answered questions to Non-Mechanics questions entered as a within-subjects variable, threat condition entered as a between-subjects variable, and age entered as a continuous covariate. The only significant effect was a main effect for children's age (F (1,17) = 5.83, p = .03, $\eta\rho^2$ = .25). A post-hoc Bivariate Correlation revealed a significant positive association between age and the proportion of children's substantive responses across all questions (r (21) = .45, p = 0.04), indicating that children provided more substantive answers as they got older. Question-type was not significant, meaning we did not observe a significant difference in children's ability to answer Mechanics of Abuse questions (M = .83, SD = .17) versus Non-Mechanics questions (M = .91, SD = .13). Threat condition also did not influence children's responding; children were equally likely to answer questions when they had been exposed to a threatening stimulus (M = .82, SD = .09) versus a non-threatening stimulus (M = .80, SD = .09).

Discussion

Previous research has uncovered that questions about the mechanics of abuse are important in legal settings. Questions about the mechanics of abuse account for almost 10% of attorneys' questions to child witnesses (Sullivan et al., 2021), and are included in 72% of forensic interviews. This is likely because these questions elicit information that is important in abuse cases, such as describing abuse encounters and differentiating between abusive and non-abusive acts. This makes it important for interviewers to understand the optimal way to deliver these questions.

Research has begun to examine the best ways to ask children about the mechanics of abuse. Research on intermediate clothing placement research has recommended using broad wh-

questions (e.g., “Where were your clothes?”) (Stolzenberg et al., 2017a; Wylie et al., 2020), but it is possible that these questions may be too vague or broad within the context of a full interview. A recent study examining how questions within the context of trial testimony found that how questions that ask for descriptions regarding the manner or outcome of an action may be too difficult for children, especially children 5 years old and younger.

The present study examined 5- to 10-year-old children’s ability to answer wh- questions about mechanics of abuse information and non-mechanics of abuse information. We had no directional hypotheses, but rather this research was exploratory in nature to examine rates of children’s ability to answer, non-how, wh- questions about the mechanics of abuse within the context of our interview questions.

Our results indicate that children generally have a high ability to answer mechanics of abuse questions. On average, children gave meaningful responses to about 80% of the questions there were asked about the mechanics of abuse. That indicates that they understood the question and did not find the questions ambiguous or unclear. This is impressive given that the mechanics of abuse information was subtle information within the context of a larger study, and there was a delay of several days between exposure to the stimulus and the memory interview. However, our data also showed that children gave non-substantive answers to an average of 16% of questions about non-mechanics of abuse questions. These questions mostly asked for descriptions of the background behind Coach Dan, which may have been more difficult for children to remember. Peterson and colleagues (1999) noted that pre-school aged children (3 to 5 years old) are more likely to remember actions than descriptions. Although our findings did not suggest misunderstanding was rampant, it does suggest that some of the questions may cause confusion and we cannot simply say asking wh- questions that are not how questions are ideal for children.

It is possible that some of our mechanics of abuse questions are clearer than others. To examine that possibility, we did a cursory review of children's responses across the different questions to see if any one question appeared more difficult than the rest. From this review, we found that there was a question that seemed more unclear compared to the other mechanics questions. More specifically, only 36% (n = 13) of children were able to accurately answer the question "What was Coach Dan Wearing?" When we reviewed children's responses, we found that instead of providing correct information, the majority of children provided a response that included either "I don't know/I don't remember" or a partially incorrect or incorrect answer that was clearly a guess based on "typical clothing" rather than a memory for "Coach Dan's" outfit. The high rates of "I don't know" could reflect that this question is particularly confusing for children and may need to be asked in a different way. Or, this could reflect a cognitive limitation with the information. More specifically, the children may not have attended to or encoded information about the specific clothing "Coach Dan" was wearing. It is also possible they encoded the information but have a difficult time recalling that type of information. All of these possibilities are supported by the findings that children are less able to answer wh- questions about descriptions than they are wh- questions about actions (Goodman et al., 1991; Peterson, et al., 1999).

We did find an effect of age regarding children's ability to answer wh- mechanics of abuse questions. Younger children performed significantly worse on these questions than did older children. This means that, although children can answer these questions, there may need to be considerations about asking these questions to young child witnesses. These findings are consistent with the findings of Henderson and colleagues (2022) who found that young witnesses (4 to 5 years old) struggled the most with "How-manner" questions. Additionally, Peterson and

colleagues (1999) noted that pre-school aged children (3 to 5 years old) are more likely to remember actions than descriptions. Together these findings imply that wh- questions about the mechanics of abuse are not ideal for use with witnesses 5 to 6 years of age and younger. However, future research is needed on why young children struggle with these questions. Since Stolzenberg and colleagues (2017a) found that children 3 to 6 years old could answer simple questions about clothing placement, it may not be that the topics themselves are too hard for children, but perhaps either the question phrasing or the context of questioning about these topics within a larger body of questions is to blame.

Additionally, both the current study and Henderson et al. (2022) included delays prior to questioning and Stolzenberg et al. (2017a) did not, so it is always possible that developmental differences in memory are driving this relationship. Research has indicated that preschool-aged children have poorer memory and less ability to understand complex conversations than older children (Ceci & Bruck, 1993).

One potential solution for young witnesses could be the use of "pairing" with asking about the mechanics of abuse. Pairing is a questioning technique whereby an interview asks a specific closed-ended question and then follows up with an open-ended response to help clarify and substantiate children's close-ended response. Prior research has found that this questioning method works well to elicit true, detailed responses from children in situations in which obtaining a response from a simple open-ended question is difficult (Stolzenberg et al., 2017b). If young children struggle with the clarity of certain mechanics questions, perhaps asking a specific closed-ended question could work to focus their attention (e.g., "Did 'Coach Dan' do something with his clothes?"), which then can be followed up with a paired open-ended question ("You said he did something with his clothes, tell me about what 'Coach Dan' did with his

clothes.”). Future research will need to be conducted with very young witnesses both to determine what aspects of these questions present a challenge for these children and to begin to evaluate potential solutions.

We did not find any significant difference in children’s ability to understand mechanics of abuse questions versus their ability to understand non-mechanics of abuse questions. Our main research goal was to determine if mechanics of abuse questions are particularly vague or unclear, not whether children are more accurate to these questions. Our results suggest that the mechanics of abuse questions are not less clear or more ambiguous than the non-mechanics questions as we did not see a significantly higher rate of unanswered mechanics of abuse questions. However, we did not directly compare the difference of accuracy scores across the two types of questions. We did not do this because the mechanics of abuse information was more salient in the stimulus than the non-mechanics of abuse questions. The non-mechanics of abuse questions were about the background stimulus and the mechanics questions were about the main character. Therefore, our stimulus was confounded with saliency. Future research will need to examine whether there are differences in children’s ability to provide *accurate* information to these questions compared to other questions. Though children’s ability to understand what information a question is asking for is important information, their ability to accurately remember and report this information is also crucial. Therefore, future research should be conducted examining children’s memory for mechanics of abuse information.

Limitations

The most significant limitation of our study is the sample size. With only 37 participants we are underpowered to find anything but very large effects. We did not observe a difference in children’s response patterns across mechanics of abuse questions and non- mechanics abuse

questions, but we cannot conclude with the current sample and data that a difference does not exist. This is because, if there is an effect, our power analysis revealed we were severely underpowered to detect it. Therefore, although this preliminary data suggests there may not be a difference across question categories, it is premature to make a strong claim about this null effect.

Another limitation to note is that our study was conducted in a controlled laboratory environment and used direct questions that did not involve abuse. Within the context of a true forensic interview, the child will be asked other questions about stimuli that are not just about the mechanics of abuse or single exposure to a specific stimulus. Although children appeared to be able to answer the majority of wh- questions in our study, it may be more difficult for them in a real-life situation when prompted with questions including information that may be sensitive or confusing.

Additionally, the majority of our wh- questions were "what" questions. Due to this, we were unable to compare children's abilities to respond to all of the different wh- questions. However, the current research and previous studies have suggested that question format is essential when interviewing children. Future research is necessary on the appropriate utilization of specific wh- questions in the context of the mechanics of abuse.

Practical Applications

Together with the previous research, this study provides support to the mounting evidence that wh- questions may be the best questions to ask when attempting to elicit mechanics of abuse information. Although results showed children were generally able to answer the wh- questions about mechanics and non-mechanics of abuse, it is too early to recommend exact questions. Additionally, due to the younger children's difficulty with these questions, they

should be used cautiously. Future research should focus on 5- to 6-year-old children and younger, as these children experience the most difficulty with mechanics of abuse questions, and they are the most vulnerable as witnesses generally.

Understanding the optimal strategies for asking children questions pertaining to the mechanics of abuse is essential. The information disclosed often contains key components of the abuse that has occurred and could affect aspects of a case such as charging or the credibility of the witness. Our results add to the emerging research on how to talk about these important topics with children, but future research must be done before any official recommendation can be made.

References

Bar-Haim, Y., Lamy, D., Lee, P., Bakermans-Kranenberg, M.J., van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. Psychological Bulletin, 133(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Benia, L. R., Hauck-Filho, N., Dillenburg, M., Stein, L. M. (2015). The NICHD investigative interview protocol: A meta-analytic review. *Journal of Child Sexual Abuse*, *24*(3), 259-279. https://doi.org/10.1080/10538712.2015.1006749

Ceci, S.J., & Bruck, M. (1993). Suggestibility of the Child Witness: A Historical Review and Synthesis. *Psychological Bulletin*, *113*(3), 403-439. https://doi.org/10.1037/0033-2909.113.3.403

Faul, F., Erdfelder, E., Lang., A-G., Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior and Research Methods*, *39*, 175-191. https://doi.org/10.3758/BF03193146

Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. Merrill-Palmer Quarterly, 37, 109–157.

Henderson, H., Sullivan, C., Wylie, B., Stolzenberg, S.N., Evans, A.D., Lyon, T.D. (March 17-20, 2022). *Child witnesses' difficulty with "how" questions: How-evaluative vs. How-manner.* This paper was presented at the American Psychology-Law Society Annual Conference, Denver, CO.

Johnson, R.R., & Aaron, J.L. (2013) Adults' Beliefs Regarding Nonverbal Cues Predictive of Violence. *Criminal Justice and Behavior*, *40*(8), 881-894. https://doi.org/10.1177/0093854813475347

Lamb, M. E., & Fauchier, A. (2001). The effects of question type on self-contradictions by

children in the course of forensic interviews. *Applied Cognitive Psychology, 15*(5), 483–491. https://doi.org/10.1002/acp.726

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, T., Yudilevitch, L. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development*, *19*(3), 627– 637. https://doi.org/10.1177/016502549601900310

Lamb, M.E., Orbach, Y., Hershkowitz, I., Esplin, P.W., Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect, 31*(11), 1201-1231. https://doi.org/ 10.1016/j.chiabu.2007.03.021

Lamb, M. E., Orbach, Y., Hershkowitz, I., Horowitz, D., Abbott, C. B. (2007). Does the type of prompt affect the accuracy of information provided by alleged victims of abuse in forensic interviews? *Applied Cognitive Psychology, 21*(9), 1117-1130. https://doi.org/10.1002/acp.1318

Lyon, T.D., McWillliams, K., Williams, S. (2019). Children as Witnesses. In N. Brewer & A. B. Douglass (Eds.), *Psychological Science and the Law* (pp. 157-181). The Guilford Press.

Lyon, T. D., Stolzenberg, S. N., McWilliams, K. (2017). Wrongful acquittals of sexual abuse. *Journal of Interpersonal Violence*, *32*(6), 805-825. https://doi.org/10.1177/0886260516657355

McWilliams, K., Williams, S., Stolzenberg, S. N., Evans, A. D., Lyon, T. D. (2021). Don't

know responding in young maltreated children: The effects of wh- questions type and enhanced interview instructions. *Law and Human Behavior, 45*(2), 124–137. https://doi.org/10.1037/lhb0000404

Myers, J. E. (1989). Allegations of child sexual abuse in custody and visitation litigation: Recommendations for improved fact finding and child protection. *Journal of Family Law*, *28*(1), 1.

Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect*, *24*(6), 733-752 https://doi.org/10.1016/S0145-2134(00)00137-X

Orbach, Y., & Lamb, M. E. (1999). Assessing the accuracy of a child's account of sexual abuse: A case study. *Child Abuse & Neglect, 23*(1), 91–98. https://doi.org/10.1016/S0145-2134(98)00114-8

Orbach, Y., & Lamb, M. E. (2001). The relationship between within-interview contradictions and eliciting interviewer utterances. *Child Abuse & Neglect, 25*(3), 323–333. https://doi.org/10.1016/S0145-2134(00)00254-4

Peterson, C., Dowden, C., Tobin., J. (1999). Interviewing Preschoolers: Comparisons Of Yes/No and Wh- Questions. *Law and Human Behavior*, *23*(5), 539-555. https://doi.org/10.1023/A:1022396112719

State vs. Emmett, 839 P.2d 781 (U.T. Sup. Ct. 1992). https://law.justia.com/cases/utah/supreme-court/1992/910077.html

Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A., Sunshine, N.

(1996). The relationship between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology*, *17*(3), 439–451. https://doi.org/10.1016/S0193-3973(96)90036-2

Sternberg, K. J., Lamb, M. E., Orbach, Y., Esplin, P. W., Mitchell, S. (2001). Use of a structured investigative protocol enhances young children's responses to free-recall prompts in the course of forensic interviews. *Journal of Applied Psychology*, *86*(5), 997–1005. https://doi.org/10.1037/0021-9010.86.5.997

Stolzenberg, S.N., & Lyon, T.D. (2014). Evidence summarized in attorneys' closing arguments predicts acquittals in criminal trials of child sexual abuse. *Child Maltreatment*, *19*(2), 119-129. https://doi.org/10.1177/1077559514539388

Stolzenberg, S.N., & Lyon, T.D. (2017). 'Where were your clothes?' Eliciting descriptions of clothing placement from children alleging sexual abuse in criminal trials and forensic interviews. *Legal and Criminological Psychology, 22*(2), 196-212. https://doi.org/10.1111/lcrp.12094

Stolzenberg, S.N., McWilliams, K., Lyon, T.D. (2017a). Spatial Language, Question Type, and Young Children's Ability to Describe Clothing: Legal and Developmental Implications. *Law and Human Behavior, 41*(3), 398-409. https://doi.org/10.1037/lhb0000237

Stolzenberg, S.N., McWilliams, K., Lyon, T.D. (2017b). The Effects of the Hypothetical Putative Confession and Negatively Valenced Yes/No Questions on Maltreated and Nonmaltreated Children's Disclosure of a Minor Transgression. *Child maltreatment*, *22*(2), 167–173. https://doi.org/10.1177/1077559516673734

Sullivan, C., St. George, S., Stolzenberg, S.N., Williams, S., Lyon, T.D. (2021) Imprecision

About Body Mechanics When Child Witnesses Are Questioned About Sexual Abuse. *Journal of Interpersonal Violence*. Advance online publication. https://doi.org/10.1177/0886260521997941

Wylie, B.E., Stolzenberg, S.N., McWilliams, K., Evans, A.D., Lyon, T.D. (2020). Young children's ability to describe intermediate clothing placement. *Child Maltreatment, 26*(1), 87-94. https://doi.org/10.1177/1077559520930825

Appendix

Coach: Clothing Placement & Physical Appearance (4)

- What color was Coach Dan's hair?
- What was Coach Dan wearing?
- What did Coach Dan wear on his face?
- What did Coach Dan do with his clothing?

Coach: Body Mechanics (7)

- What mood do you think Coach Dan was in?
- You said he was X (restate child's answer to the previous question), what made him seem X?
- What was the expression on Coach Dan's face?
- What was he doing with his body?
- What was he doing with his hands?
- What was the tone of his voice?
- Was Coach Dan close or far from the screen?

Speech: Background (7)

- What was in the picture behind coach dan?
- What were the people doing?
- How many people were there?
- Were they just boys, just girls, or both?
- What color was the room?
- What mood do you think the characters in the picture were in?
- You said they were X (restate child's answer to the previous question), what made them seem X?