11. Twenty-five years of interviewing research and practice: Dolls, diagrams, and the dynamics of abuse disclosure.

Thomas D. Lyon, University of Southern California
Twenty-five Years of Interviewing Research and Practice: Dolls, Diagrams, and the Dynamics of Abuse Disclosure

Thomas D. Lyon, JD, PhD

A great deal of research in the past 25 years has contributed to our understanding of how best to interview children about suspected maltreatment. The disastrous failures of the highly publicized day care abuse cases led to a flood of research, initially emphasizing the failures of conventional approaches, and more recently highlighting the potential for eliciting complete and accurate reports. If a child has disclosed abuse, and is willing to disclose again, we know what to do. Research supports the use of interview instructions, narrative practice rapport building, and the use of open-ended questions to elicit and to elaborate on the child’s report (Saywitz et al., 2011). These elements are found in an increasing number of interview protocols, most notably the NICHD structured protocol (Lamb et al., 2008).

However, prior disclosure is a big if. The likelihood that abused children will refuse to acknowledge abuse has long been recognized (Pollack, 1909). The problem of reluctance is recognized by proponents of the NICHD protocol (Lyon et al., 2009), and researchers continue to seek means of overcoming reluctance through improvements (Hershkowitz, 2011). It is fair to say that whereas the focus since the 1980s has been on reducing false allegations, researchers have increasingly turned to means of increasing true allegations.

Nevertheless, tensions within the field exist among both interviewers and researchers regarding the best next steps for interviewing. An enduring debate that nicely captures these tensions concerns the use of anatomical dolls and diagrams. Dolls were developed in the 1970s (Koocher et al., 1995), and their use was widespread in many jurisdictions by the mid-1980s (Boat & Everson, 1988; In re Rinesmith, 1985). In the 1980s, Groth (1984) developed anatomically detailed diagrams of children for use in sexual abuse interviews. The theory was that young children might better describe their abuse through use of the dolls and diagrams, overcoming developmental and motivational difficulties in disclosing.

When APSAC was founded in 1987, researchers had only just begun to examine anatomical dolls. The first study examining the use of dolls in interviewing was published in 1986 (White et al., 1986), and the results were reassuring: Children with other evidence of abuse responded differently to questioning than children for whom there were no suspicions. However, studies observing free play had raised red flags regarding interpretation of children’s free play with the dolls (Gabriel, 1985; Jampole & Weber, 1987), and experts had made questionable interpretations of behaviors, such as digital insertion (In re Cheryl H., 1994), that were later found to be quite common among children who played with the dolls (Cohn, 1991). Diagrams received less attention; their use was not systematically studied until the 1990s (Steward et al., 1996).

Twenty-five years later, the disagreements continue. Although observations of children’s free play with the dolls have fallen out of vogue, the use of dolls and diagrams to elicit disclosures or to clarify reports is still popular. On the one hand, many interviewers support their use (Anderson et al., 2010; Hlavka et al., 2010), and Faller (2005, 2007) reviewed the research favorably. On the other hand, the experimental work published since 2000 has been uniformly critical (Brown et al., 2007; Bruck, 2009; Bruck et al., 1995, 2000; Otgaar et al., in press; Poole & Dickinson, 2011; Wilcock et al., 2006), and most research reviews have been similarly negative (Brown, 2011; Dickinson et al., 2005; Pipe & Salmon, 2009; Poole et al., 2011; Salmon, 2001).

The difference is attributable to unspoken value judgments and a lacking appreciation of the dynamics of sexual abuse disclosure. It is not enough to prove that dolls and diagrams elicit more details; one must have some means of determining whether those details are true. It is also insufficient to limit one’s focus to false allegations: one must always weigh the costs against the potential benefits. The best studies examine genital touch in medical contexts, because this provides the closest analog to sexual touch, enables one to assess accuracy, and allows one to assess the effects of dolls and diagrams on children who have and have not been touched. This research warns against doll and diagram use in very young children and counsels caution in their use with older children. But the risks have been exaggerated by some research, and reasonable minds still disagree about the potential utility of dolls and diagrams when nondirect questions fail to elicit disclosures.
Field Research on Dolls and Diagrams

Interviewers are likely to be most impressed with results in the field, because field research has the advantage of external validity; these are real cases of alleged sexual abuse. Ironically, that is also its disadvantage; the accuracy of any additional details elicited by dolls or drawings often cannot be ascertained. Some field research appears to provide support for dolls and diagrams; early doll studies were positive (e.g., Leventhal et al., 1989; White et al., 1986), and two recent studies utilizing body diagrams (with the genitalia obscured) found that diagrams elicited new details when introduced at the end of an interview (Aldridge et al., 2004; Teoh et al., 2010). However, two recent studies examining doll interviews found that they were no more productive than interviews without dolls (Lamb et al., 1996; Thierry et al., 2005).

Skeptics can discount either positive or negative findings. The studies examining non-NICHD interviews can be criticized for the failure of the interviewers to utilize all available means of eliciting complete reports through open-ended questions. When additional details are elicited, it is often not clear if the dolls or diagrams are responsible, unless the researchers compared introduction of dolls and diagrams with a separate condition in which children are simply asked to recall the abuse a second time (Salmon et al., 2011). Furthermore, studies finding the elicitation of additional details do not tell us whether dolls or diagrams are useful in eliciting disclosures from children who fail to disclose in response to other prompts.

When additional details are not elicited, it is sometimes questionable whether children who received and did not receive the dolls or diagrams are comparable. For example, if interviewers chose when and whether to use dolls, they may have selectively done so with less productive children, which would make dolls and diagrams look less productive than they really are (Faller, 2005). In technical terms, unless children are randomly assigned to the doll condition and the no-doll condition, one doesn’t know what to make of any differences. Notably, these methodological difficulties can be overcome with more carefully controlled research, but field research is extremely difficult: It is very difficult to obtain the necessary consents, and it is time-consuming and expensive (e.g., one has to train and carefully monitor the interviewers). Furthermore, even if other methodological concerns are met, the accuracy issue almost always remains, because of the lack of clear corroborative evidence in most abuse cases.

Medical Exam Studies With Dolls and Diagrams

It thus seems likely that the debates over the utility of dolls must look to experimental evidence, but then the question is: What is the appropriate analog to sexual abuse? The initial wave of research on dolls and diagrams turned to medical examinations, which had a number of advantages.

First, there could be conditions in which children either had or had not experienced genital touch. This allows one to calculate both true positive and false positive rates for any technique, which in turn enables one to assess the probative value of a disclosure under different circumstances. It is essential to be able to calculate both rates. Imagine a study that included only children who had been touched. A method might increase disclosures, but be essentially worthless if it increased false disclosures by the same amount. But we wouldn’t know its effects on false disclosures if all the children had been touched. But by the same token, any study that includes only children who haven’t been touched is equally incomplete. A method might increase false disclosures, but be valuable if it increased true disclosures by a much larger amount.

Second, medical examinations specifically enable one to inquire into genital touch. Researchers examining genital touch in medical examinations understood that genital touch is different than other types of touch; once children are out of diapers, toilettained, and can bathe themselves, their genitals are less likely to be touched by adults as part of caretaking. When they are touched on their genitals, and particularly if the touch is invasive, they are likely to experience it as unpleasant: It is salient, often embarrassing, and sometimes disgusting to the child. An obvious limitation to studying medical examinations is that when a doctor touches a child’s genitalia, he or she does so for a valid medical purpose. The child’s parent is likely to be present, and the doctor will not say or do anything to suggest that the touching is secretive or in some way wrong. Hence, the dynamics of sexually abusive genital touch are likely to be perceived as more wrongful and more embarrassing.

A series of studies examined children’s reports of genital touching as part of well-child examinations, and the findings can be summarized quite easily (Bruck et al., 1995, 2000; Saywitz et al., 1991; Steward et al., 1996). When 3–7-year-old children are asked free recall questions about the medical examination, they only rarely disclose genital touch if they have been touched, and never disclose genital touch if they haven’t been touched (Saywitz et al., 1991; Steward et al., 1996). When the interviewer moves to direct questions utilizing a doll or drawing, the likelihood that children disclose touch increases. Bruck and colleagues (1995, 2000) found that 2–4-year-old children exhibited relatively low true positive rates (only about 50% of those who were touched said that they had been touched), and high false positive rates (about 50% of those who were not touched claimed that they had). The fact that the true positive and false positive rates were almost equal suggested that children were responding randomly, and it meant that a disclosure of touch in response to a direct question was not probative. Steward and colleagues (1996) found that 3–6-year-old children exhibited higher true positive rates than false positive rates, such that a disclosure of touching was weak to moderate evidence of touching. Saywitz and colleagues
Subsequent reviews of the literature have evaluated these studies differently. Faller (2005) interprets Steward’s and colleagues’ study as supporting doll and diagram use; whereas, others (e.g., Poole & Dickinson, 2011) emphasize the increase in false positives. There may be implicit value judgments being made here (Ceci & Friedman, 2000). Is the increase in error justifiable, given the increase in true disclosures? When Poole and Dickinson (2011) conclude that diagrams should not be used to elicit disclosures unless there is evidence of abuse akin to “images or a definitive medical finding,” they are assuming that a conclusion that abuse occurred requires an extremely high standard of proof (p. 668). I suspect that Faller would put more emphasis on the need to utilize sensitive measures in order to avoid missing true cases. Of course, other considerations must come into play. Is this a criminal case? Is the child potentially at risk of further abuse?

The four medical examination studies leave a number of questions unanswered. None of the studies considered whether improved methods of eliciting free recall might increase true disclosures; none utilized narrative practice rapport building before recall or cued invitations after. None provided more than brief follow-up to disclosures to determine whether elaboration might make true and false disclosures distinguishable. Finally, none compared questioning with a doll or diagram with direct questions about genital touch alone. It might be the case that dolls and diagrams are less necessary with improved interviewing, because children who are more comfortable with the interviewer and more talkative are more likely to disclose. However, for children who fail to disclose despite improved interviewing, dolls and diagrams might be less dangerous to use as a backup, because the accuracy of the disclosures can be tested through testing the child’s ability to elaborate on an acknowledgement of touch.

Recent Research on Dolls and Diagrams: No More Genital Touching

Sadly, these questions have remained unanswered, in part because the research conducted in the last decade has failed to utilize the medical examination paradigm. Instead, children experience nongenital touch and are asked questions with the assistance of diagrams that fail to depict the genitalia. The touching that children experience is typically not very salient and certainly not embarrassing. The fact that none of the children experience genital touch means that the true positive rate of the diagrams cannot be determined—the research can calculate only the rate of false allegations. The fact that the touch is not embarrassing means that there is no reason to assume that children are reluctant to disclose touch. Despite these limitations, the researchers often conclude in very strong terms that the diagrams are not useful to questioning children about suspected sexual abuse (e.g., Willcock et al., 2006).

In the subsequent studies, body diagrams led to some false reports of touch, though not always claims of genital touch. Willcock and colleagues (2006) found that one month after interacting with a man who touched them in five innocent places, 11% of 5–6-year-olds disclosed genital touch when questioned with a clothed body diagram. Brown and colleagues (2007) found that 4 to 6 weeks after experiencing seven innocent touches (e.g., tickling the feet, squeezing the wrist), 4% of 5–7-year-olds disclosed genital touch
either when directly asked or when questioned with an unclothed body diagram. In free recall, none of the children mentioned any touching, and a large percentage failed to report touching in response to direct questions (with or without the diagrams). Poole and Dickinson (2011) found that 4 months after being touched on the wrist (a wrist band) and shoulder (a sticker), 0% of 4-9-year-olds disclosed genital touch when questioned with an unclothed body diagram. Without the diagrams, only 1% of children mentioned the touch that did occur; in the diagram condition, 8% did so. Poole and Dickinson also included a group in which children received suggestions from their parents that they did in fact receive touches that they had not experienced; the rates of false reports of those touches were similar in the no-diagram and diagram conditions. In all three studies, children falsely reported other types of touch, although Brown and colleagues (2007) found similar rates regardless of whether a direct question was asked or the body diagrams were used.

Unfortunately, the studies are not terribly useful in helping us assess the potential utility of diagrams in questioning children about genital touch. Because there was no condition in which children were touched on their genitalia, one cannot calculate the percentage of children who were touched who revealed with or without the diagrams. Children often showed very low rates of touch disclosure, but there is no reason to assume that children were reluctant to disclose any of the touches that occurred. It is more likely that they simply forgot the touching or found it unremarkable. (Cf. Bruck, 2009, who found that children underreported touching immediately after a staged event.)

With respect to the false reports of genital touch, the problem is that the diagrams omitted the genitalia. Poole and Dickinson (2011) assert that this explains why they didn't obtain any false reports of genital touch, but they provide no support for their apparent belief that explicit depiction would increase the likelihood of error. Rather, the opposite problem might be at work: When the genitalia are not depicted, this increases the risk of misunderstanding. This is a possible explanation for the higher rate of false reports of genital touch in Willcock and colleagues' study, in which clothed diagrams were used; Brown and colleagues suggest that the clothed drawings made it "more difficult to specify where touches occurred" (Brown et al., 2007, p. 40). Indeed, a recent study comparing clothed with unclothed diagrams found that younger children produced more accurate details in response to the unclothed diagrams, and it speculated that the lack of clothing facilitated children's recognition of the parts of the body (Otgaar et al., in press).

The Importance of Studying Reluctant Disclosure
Despite these limitations, the researchers conclude that body diagrams are not useful for eliciting reports of genital touch in sexual abuse investigations. With respect to the argument that one ought to be studying touch that is analogous to sexual abuse—probably genital touch, at least touch that children find embarrassing—they make different arguments. Poole and Dickinson (2011) acknowledge that abuse may be "embarrassing or traumatic," but argue that this is irrelevant if one is interested in examining false allegations, because those involve children who have not been abused (p. 668). This argument misses two points.

First, embarrassment affects not only children who have been touched but also children who haven't. Children who are aware that genital touch is unusual and embarrassing will be less inclined to false alarm to suggestions of genital touch than to suggestions of innocuous touch. Steward and colleagues, for example, found that whereas 42% of children falsely reported touches to the ears in free recall 6 months after a medical exam, 0% falsely recalled genital touch.

Second, false allegations cannot be assessed in a vacuum, unless one adopts the value judgment that any increase in false positives is unacceptable. The question is always how often children who were touched disclose compared with how often children who weren't touched false alarm. (Even with respect to the touches they did study, Poole and Dickinson (2011) couldn't provide this analysis: Children who were touched and children who weren't touched were not comparable, because all of the children who weren't touched were subjected to repeated suggestions of touching before being questioned.)

Salmon and colleagues (2011) make the point that that reports of incidental touch may be relevant because abuse "in the early phases" is often initiated through purportedly accidental touch during daily activities (General Discussion section, para. 3). This is true, and their results (which concern the ability of diagrams to clarify reports of touching) suggest that diagrams don't facilitate disclosure of touching that the child found unremarkable. It is not clear, however, why this is a disadvantage; it would be dangerous to characterize touch as abusive if it could have been accidental. Furthermore, the argument implicitly recognizes that if one is investigating overt abuse that is recognized as such by the child being questioned, studies examining children's reports of incidental touching are less relevant.

Most remarkably, some researchers argue that sexually abused children are not reluctant to disclose. Bruck and Ceci (2009), for example, discuss a study in which they found that large percentages of children were initially reluctant to acknowledge misbehavior at school. Asserting that the study has no relevance for understanding possible denial of sexual abuse, they explained that "the motives to deny an actual punishment are quite different from denying sexual abuse. The former involve protecting oneself from revealing an embarrassing wrongdoing" (p. 158). Ironically, this is a concise description of how many (if not most) abused children perceive abuse: an embarrassing wrongdoing. Indeed, when adult survey respondents are asked why they never disclosed
abuse as a child, two of the most common reasons are embarrassment and a fear that they would be blamed for the abuse (Anderson et al., 1993; Fleming, 1997).

Until researchers acknowledge the importance of understanding the dynamics of sexual abuse disclosure, their research will have limited applicability to abuse investigation. Even if they cannot obtain permission to study contexts in which genital touch occurs, they should take account of motivational barriers to disclosure, and design their studies accordingly. Indeed, fear of punishment provides a promising laboratory analog to disclosure reluctance. Lab studies examining children's concealment of transgressions, and means of encouraging them to disclose, have revealed the advantages of eliciting a promise to tell the truth and the limited advantages of reassurance (Evans & Lee, 2010; Lyon & Dorado, 2008; Lyon et al., 2008; Talwar et al., 2002, 2004).

Conclusion: Dolls, Diagrams, and the Future

Ultimately, I don't have any easy answers for practitioners who are considering whether to use dolls and diagrams in questioning children about abuse. My personal view is that they should be used only as a last resort and avoided altogether with children under 4 years of age. After one has worked through the disclosure questions that are provided by the NICHD protocol, then direct questions about genital touch could be used with caution and only when a subsequent interview is not practical (or when delay may endanger the child). But I would stress that my view is based on limited knowledge, on value judgments, and primarily on the research that best applies: studies examining children's true and false reports of genital touch.

Child interviewing research has enabled interviewers to make great strides in the past 25 years. We can now point to evidence-based approaches to interviewing that increase the productivity of children's reports without increasing the likelihood of false reports. The next step is to identify the best means of eliciting disclosures from children who are reluctant to reveal abuse, whether because of fear, embarrassment, guilt, shame, or other motivational barriers. The most progress will be made if researchers learn as much as they can about the dynamics of sexual abuse and the principles of cognitive, social, and language development and if they are as open as they can be about the methodological strengths and weaknesses of their work. Moreover, we must all be mindful of the devastation wreaked by both false allegations and false denials of abuse.

References


---

**About the Author**

Thomas D. Lyon, JD, PhD, is the Judge Edward J. and Ruel L. Guirado Chair in Law and Psychology at the University of Southern California Gould School of Law. A magna cum laude graduate of Dartmouth College and Harvard Law School, Professor Lyon received his PhD in developmental psychology from Stanford University. He was an attorney for the Children's Services Division of the Los Angeles County Counsel and a research associate at Harbor-UCLA Medical Center prior to joining USC Law in 1995. Professor Lyon is past president of the American Psychological Association's Section on Child Maltreatment (Division 37) and a former member of the Board of Directors of the American Professional Society on the Abuse of Children. He has published more than 50 papers in law reviews, psychology journals, and books; has authored or coauthored more than 80 research presentations at psychology and law conferences; and has conducted more than 170 trainings with judges, attorneys, law professors, social workers, psychologists, and reporters. Preparation of this article was supported by NICHD grant HD047290. Contact: tlyon@law.usc.edu