Fall October 21, 2020

80. Understanding expert testimony on child sexual abuse denial after New Jersey v. J.L.G.: Ground truth, disclosure suspicion bias, and disclosure substantiation bias.

Thomas D. Lyon
Shanna Williams, *McGill University*
Stacia N Stolzenberg, *Arizona State University*

Received: 8 October 2020 | Revised: 30 September 2020 | Accepted: 21 October 2020

DOI: 10.1002/bsl.2490

WILEY

SPECIAL ISSUE ARTICLE

Understanding expert testimony on child sexual abuse denial after *New Jersey v. J.L.G.*: Ground truth, disclosure suspicion bias, and disclosure substantiation bias

Thomas D. Lyon[1] | **Shanna Williams**[2] | **Stacia N. Stolzenberg**[3]

[1]University of Southern California Gould School of Law, Los Angeles, CA, USA

[2]McGill University, Montreal, Canada

[3]Arizona State University, Tempe, AZ, USA

Correspondence
Thomas D. Lyon, USC Gould School of Law, 699 Exposition Blvd, Los Angeles, CA 90089-0071, USA.
Email: tlyon@law.usc.edu

[Corrections made on 23 November 2020, after first online publication: Tables 1 and 4 and an in-text citation in Section 6 have been updated in this version.]

Abstract

The New Jersey Supreme Court held in *New Jersey v. J.L.G.* (2018) that experts can no longer explain to juries why sexually abused children might deny abuse. The court was influenced by expert testimony that "methodologically superior" studies find lower rates of denial. Examining the studies in detail, we argue that the expert testimony was flawed due to three problems with using child disclosure studies to estimate the likelihood that abused children are reluctant to disclose abuse: the ground truth problem, disclosure suspicion bias, and disclosure substantiation bias. Research identifying groups of children whose abuse can be proven without reliance on disclosure reveals that denial of sexual abuse is common among abused children.

1 | INTRODUCTION

In *New Jersey v. J.L.G.* (2018), the New Jersey Supreme Court re-evaluated the admissibility of expert testimony on child sexual abuse accommodation syndrome (CSAAS). CSAAS describes sexually abused children's secrecy, helplessness, entrapment, delayed disclosure, and retraction (Summit, 1983). The court held that, although scientific evidence supports the proposition that child sexual abuse victims delay disclosure, the other elements of CSAAS are not adequately supported and are not allowed in expert testimony. Specifically, the court found that experts could not explain to juries why sexually abused children might deny abuse. The court was influenced by reviews of the research asserting that "methodologically superior" studies find lower rates of denial.

This paper examines CSAAS and briefly reviews its reception by the courts and by researchers. We then examine the New Jersey Supreme Court case in *J.L.G.* and the research it found most persuasive. We argue that the

court was misled, and outline the methodological challenges in examining disclosure rates among children questioned about sexual abuse. In particular, we will show how problems with establishing ground truth and biases in sample selection lead one to overlook substantial evidence of denial.

1.1 | CSAAS: The original article and acceptance by the courts

Summit (1983) discussed how repeated sexual abuse by an adult in a position of trust is initiated and maintained in secret. His article described how sexual abuse progresses through exploitation of the helpless and dependent child, and through inducements to keep the abuse a secret. It described how the child's inability to report the first acts of abuse guarantees future victimization and leads the child to blame herself for the abuse. Disclosure, if and when it occurs, is delayed and unconvincing, due to the child's ambivalence about the utility of telling and the reluctance of the non-offending parent to believe the child. Finally, abused children may recant their allegations in response to the negative consequences of disclosure, most notably the rejection by those to whom they turn for support. Summit (1983) cited data showing that most child sexual abuse is perpetrated by trusted adults rather than strangers (Finkelhor, 1979, 1980; Gagnon, 1965; Groth, 1978; Russell, 1983), that most abuse victims delay disclosing, if they disclose at all (Finkelhor, 1979, 1980; Gagnon, 1965; Russell, 1983), and that victims' disclosures often elicit unsupportive reactions from adults (Finkelhor, 1980).

CSAAS is subject to criticism on a number of grounds, particularly if its purpose in trials is misunderstood. First, the term "syndrome" is potentially misleading. It suggests that the elements constitute symptoms of sexual abuse, and that if a child exhibits these symptoms this proves that sexual abuse occurred. Indeed, Summit (1992) subsequently regretted using the term "syndrome." Moreover, one should not expect that sexually abused children exhibit all of the behaviors described by CSAAS. For example, a child may delay reporting, yet never recant. Second, CSAAS was based only in part on research. Summit (1983) noted that many of his arguments were based on his work as a clinical consultant and endorsements from professionals, victims, and their families. He did not attempt to exhaustively review the research available at the time. A number of commentators have subsequently criticized CSAAS on these and other grounds (see, e.g., O'Donohue & Benuto, 2012).

Expert testimony describing CSAAS became a popular tool among prosecutors in child sexual abuse cases. The New Jersey Supreme Court approved expert testimony regarding CSAAS in *New Jersey v. J.Q.* (1993). Because juries could misunderstand expert testimony, however, subsequent cases limited the purpose and scope of CSAAS testimony. The court repeatedly emphasized that accommodation testimony is not proof that abuse occurred (*New Jersey v. R.B.*, 2005, p. 308 ("Testimony concerning [accommodation] is not admissible as substantive proof of child abuse"); *New Jersey v. J.R.*, 2017). Rather, it fulfils an "exclusively rehabilitative role" (*New Jersey v. J.R.*, 2017, p. 417). That is, experts do not testify that if a child denies abuse, or exhibits other CSAAS characteristics, this is evidence that the child was abused. Rather, the experts' testimony rebuts the assumption that denial of abuse proves that abuse did not occur. Other states took a similar approach. Most state courts considering the issue have held that CSAAS testimony is admissible as rehabilitative evidence (*New York v. Spicola*, 2011), but not as proof of abuse (see, e.g., *Connecticut v. Favoccia*, 2012).

The jury instructions required in New Jersey in cases in which CSAAS testimony had been allowed illustrate the limited purpose of the testimony. Juries were told that CSAAS "relates only to a pattern of behavior of the victim which may be present in some child sexual abuse cases," and that the testimony was admitted "only to explain that the behavior of the alleged victim was not necessarily inconsistent with sexual abuse." They were told that CSAAS "may help explain why a sexually abused child may delay reporting and/or recant allegations of sexual abuse and/or deny that any sexual abuse occurred" (Model Jury Charge (Criminal), Child Sexual Abuse Accommodation Syndrome (rev. 5/16/2011)).

1.2 | Subsequent research and *New Jersey v. J.L.G.*

A substantial amount of research examining issues relevant to child sexual abuse disclosure has appeared in the years since Summit (1983), including research examining children's disclosures when questioned about abuse (e.g., Hershkowitz, Lamb, & Katz, 2014), surveys of adults asking whether they had ever disclosed abuse (e.g., McGuire & London, 2020), experimental work examining children's disclosures of wrongdoing (e.g., Williams, McWilliams, & Lyon, 2020), and observational research examining the modus operandi of child sexual abuse perpetrators (e.g., Leclerc, Proulx, & Beauregard, 2009). Most relevant to this paper, reviews of adult surveys and children's disclosures have supported some of Summit's claims (Alaggia, Collin-Vézina, & Lateef, 2019; Lemaigre, Taylor, & Gittoes, 2017), though two of the most-cited reviews only partially support CSAAS (London, Bruck, Ceci, & Shuman, 2005; London, Bruck, Wright, & Ceci, 2008). London and colleagues found that adult surveys support the propositions that sexual abuse victims usually delay disclosing abuse (London et al., 2005, 2008), and that delays are longer the closer the relationship of the child to the perpetrator (London et al., 2008), but concluded that studies examining children's disclosures reveal very low rates of denial. In other words, they argued that children will delay disclosing abuse, but only rarely deny abuse when directly asked.

In reversing the long-standing acceptance of CSAAS testimony in New Jersey courts, the New Jersey Supreme Court in *J.L.G.* relied heavily on the reviews by London et al. (2005, 2008), which were co-authored by one of the defense experts, Dr Maggie Bruck. With respect to denial of abuse, the court noted that "In [Dr Bruck's] judgment, weaker studies—that involved children later found to have made false allegations or children who may have been subjected to suggestive techniques, for example—produced higher rates of denial. Studies with better methodologies produced low rates." (p. 462). In her expert report, Dr Bruck elaborated on this point. Citing the review papers, Bruck reported that they found that "the methodological adequacy and sampling approach of each study was directly related to the denial and recantation rates observed" (Bruck, 2017, p. 4).

> [F]or the six methodologically superior studies, the average rate of denial was only 14%. The average rate of recantation was also low for this last set of studies, 7%. These six studies provided denial and recantation data on children whose abuse status was considered "highly probable" (i.e., valid) and who were not selected because of special characteristics (e.g., sexually transmitted infections with no prior disclosure; peer abuse; single parent families). These latter studies of sexually abused children's response patterns indicate that if they are directly asked, most children do not deny their abuse, but disclose abuse. Also, a small minority of these children will recant their allegations (Bruck, 2017, pp. 4–5).

The six studies are described in the 2008 review paper (London et al., 2008), and five of them are similarly described in the 2005 review paper (London et al., 2005). The claim that "methodologically superior" studies document low rates of denial has become a recurrent claim in subsequent reviews (Bruck & Ceci, 2004, 2009; Bruck, Ceci, & Principe, 2007; Ceci & Bruck, 2013; Ceci, Kulkofsky, Klemfuss, Sweeney, & Bruck, 2007; London & Kulfofsky, 2009), as well as in recent commentary on expert testimony (Zajac, Garry, London, Goodyear-Smith, & Hayne, 2013). Because the claim presents a major challenge to key aspects of CSAAS testimony, it is worth exploring in depth.

2 | METHODOLOGICAL CHALLENGES: GROUND TRUTH, SUSPICION BIAS, AND SUBSTANTIATION BIAS

There are three problems with using studies of child forensic interviews to estimate the likelihood that a truly abused child will deny or recant abuse: the ground truth problem, disclosure suspicion bias, and disclosure substantiation bias. All three problems stem from relying on disclosure as the primary evidence of abuse.

The first problem is the ground truth problem: one is often unsure whether children have in fact been abused. One cannot simply look at the disclosure rates among children suspected of being abused, because suspicions may be unfounded, and disclosures may be untrue. A low rate of disclosure could simply reflect a high rate of unfounded suspicions. A high rate of disclosure could simply reflect a high rate of false disclosures.

The second problem is disclosure suspicion bias. Disclosure suspicion bias occurs when disclosure is the reason abuse is suspected. If disclosure increases suspicions of abuse, the percentage of children disclosing abuse in samples suspected of having been sexually abused will be inflated. Suspicion bias may operate in at least two ways. If an abused child never discloses abuse, this will decrease the likelihood that anyone will question the child about abuse. If an abused child shows soft signs of abuse (e.g. sexualized behavior), but does not acknowledge abuse to caretakers, this will likely decrease the likelihood that anyone will have the child formally evaluated. Children who never tell, or who deny abuse when questioned by caretakers, may, as a result, be disproportionately excluded from samples of children evaluated for suspected sexual abuse.

Based on representative population surveys of adults, we know that most victims of child sexual abuse do not disclose abuse during childhood, and only 10% are known to authorities (Lyon, 2009). Hence, a representative sample of abused children would contain a small percentage who had previously disclosed. If samples of children assessed for sexual abuse contain high percentages of children who disclosed prior to the assessment, then this is evidence of suspicion bias.

The third problem is disclosure substantiation bias. Disclosure substantiation bias occurs when disclosure is a reason why abuse is substantiated by authorities. If disclosure increases the likelihood that abuse will be substantiated, then the percentage of disclosure in substantiated samples of abuse will be inflated. If disclosure is the *only* way in which abuse can be substantiated, then substantiated cases of sexual abuse will have 100% disclosure rates.

In *J.L.G.*, the focus was on the ground truth problem. The court relied on Bruck's claim that the "methodologically superior studies" examining "valid" cases of abuse found higher rates of disclosure. However, neither the court nor Bruck discussed the extent to which the studies determined validity on the basis of a disclosure from the child, which fails to solve the ground truth problem. Furthermore, although disclosure suspicion bias and disclosure substantiation bias have been discussed in the literature (Azzopardi, Eirich, Rash, MacDonald, & Madigan, 2019; Lyon, 2007; Rush, Lyon, Ahern, & Quas, 2014), and acknowledged to some extent by Bruck and her co-authors in their published work (London et al., 2008), they were not discussed in *J.L.G.*, and only briefly mentioned by Bruck.

2.1 | Evidence of abuse independent of disclosure

A step toward avoiding the methodological problems is to focus on those cases for which there is evidence of abuse independent of disclosure. If one can be more confident that children classified as abused were in fact abused, the ground truth problem is reduced. If independent evidence allows one to substantiate abuse without a disclosure, then substantiation bias is reduced as well. Under some circumstances, independent evidence can even reduce suspicion bias. Recall that suspicion bias occurs if abuse is initially suspected because of the disclosure of abuse. If independent evidence is the *first indication* that a child has been abused, then suspicion bias is less of a concern.

It is important to add, however, that simply identifying cases with non-disclosure corroborative evidence of abuse is not enough. If disclosure increases the likelihood that corroborative evidence will be discovered, or if corroborative evidence increases the likelihood that disclosure will occur, then estimates of disclosure in corroborated cases of sexual abuse will be inflated. For example, criminal convictions are often characterized as corroborating abuse (DeVoe & Faller, 1999), but are highly dependent upon disclosure. The less forthcoming and less consistent child witnesses are less likely to be referred for prosecution by the police (Stroud, Martens, & Barker, 2000) and more likely to be rejected for prosecution by prosecutors (Gray, 1993). Subsequent

inconsistencies or recantations increase the likelihood of dismissals (Gray, 1993). Hence, disclosure rates associated with convictions are likely to be inflated (indeed, they are virtually 100%; Faller & Henry, 2000).

Confessions are corroborative evidence of abuse. However, confessions are not clearly independent of disclosure. Confessions both trigger disclosure and are triggered by disclosure. If the child has disclosed, this can be a tool to elicit confessions, and if the offender has confessed, this can be a tool to elicit disclosures. This positive relation will inflate disclosure rates in cases with confessions.

Medical evidence of abuse may have fewer dependence problems, but such problems are nevertheless likely to occur. The fact that a child exhibits medical signs of sexual abuse may lead interviewers to push harder to elicit a disclosure. If this does, in fact, increase the likelihood of disclosure, then the percentage of disclosures among cases with medical evidence will be inflated. Conversely, the fact that a child has disclosed abuse may make medical examiners look harder for medical signs of abuse, or may lead them to call ambiguous medical conditions supportive of abuse. To the extent that this increases the likelihood that positive medical evidence will be found, the percentage of disclosures among cases with medical evidence will be inflated.

Suspicion bias and substantiation bias will be minimized to the extent that the medical condition is both reliably diagnosed and diagnosed without knowledge of the child's disclosure. Examining research on disclosure rates among children with medical evidence of abuse is thus likely to reduce but not eliminate substantiation bias and suspicion bias.

2.2 | Goals of this paper

This paper will discuss the findings of the six "methodologically superior studies" that the experts in *J.L.G.* argued proved that denial is rare. The experts' analysis of the studies accentuated substantiation bias and did not solve the ground truth problem. The experts examined the subset of cases in which abuse could be substantiated. However, because substantiation was dependent upon disclosure, this inflated disclosure rates. Because disclosures might be false, this failed to solve the ground truth problem. The studies do not support the interpretation that the more certain one is that abuse occurred, the more likely it is that a child will disclose abuse. Indeed, the authors of the studies themselves tended to avoid this reasoning.

For each study, we will first describe the purpose of the study. We will then discuss the overall and substantiated case disclosure rates, and show how they reflect substantiation bias. We will then discuss whether the study offers a means out of the substantiation bias and ground truth problems by identifying subgroups with independent evidence of abuse. Finally, we will discuss additional insights from the studies regarding reluctance and denial that are obscured by substantiated case disclosure rates. For example, we will discuss the prior disclosure problem, which concerns cases in which children disclose abuse, but then recant abuse during the formal evaluation and thus cannot be substantiated as abused.

3 | DUBOWITZ, BLACK, AND HARRINGTON (1992)

The purpose of the study was to identify the contribution that different aspects of the sexual abuse evaluation process make to the diagnosis of abuse, including interviews with children and medical examinations. The study included 132 children under 12 years of age ($M = 6$) evaluated for suspected sexual abuse. The multidisciplinary team's assessment of the likelihood of abuse was coded as low, moderate, or high. Children's disclosure was coded as none, partial, or full. The results of the medical examination could be normal, compatible with abuse, or indicative of abuse. The introduction noted that "experts in the field of child maltreatment agree that the history obtained from the child is usually the most important evidence in diagnosing sexual abuse" (p. 688). The authors concluded that "both the child's disclosure of abuse and the physical examination findings significantly contributed to the

team's assessment of abuse; without either kind of evidence, the team was unlikely to conclude that the child had been abused" (p. 691). Hence, substantiation was highly dependent on disclosure unless there was medical evidence of abuse.

The study reported that, of the 99 children with complete information, 32% gave a clear disclosure, and 27% a partial disclosure. Based on this and other data reported by Dubowitz et al., London et al. (2005, 2008) calculated an overall disclosure rate of 59%, and a substantiated disclosure rate of 83%. In a table (Dubowitz et al., 1992, Table 3, p. 691), Dubowitz and colleagues reported the relation among the assessment likelihood of abuse, disclosure, and medical evidence of abuse. The numbers in the table are reproduced here (Table 1), and numbers calculated from the table are in brackets. Based on the numbers, one can construct both the 59% disclosure rate ((1d+2d+4d+5d)/99) and what London and colleagues identify as the substantiated disclosure rate of 83% ((1d+2d)/(1d+2d+3d)) (London et al., 2005, Table 2 pp. 206, 211; London et al., 2008, Table 2, p. 36). Based on these numbers, London and colleagues noted that the greater the certainty of abuse, the higher the disclosure rate. However, this says nothing about the true disclosure rate; it merely illustrates the dependence of substantiation on disclosure.

Perhaps the best explanation of the circularity of London and colleagues' interpretation comes from London and colleagues themselves. In their 2005 paper, London and colleagues noted that "in terms of the studies that are included in this article, the children's statements at the time of formal interview are used as indicators of abuse. But this is a circular exercise whereby children who make spontaneous disclosures with much elaboration, for example, are categorized in the 'high-certainty' abuse group" (p. 210). (cf. London et al., 2008, p. 39: "using children's statements as indicators of abuse to some extent is tautological in studies of disclosure.")

Dubowitz and colleagues did not fall prey to this circularity. The authors never argued that 59% is the disclosure rate among abused children, recognizing that the non-disclosers may not have been abused. Nor did they assert that 83% is the actual disclosure rate; indeed the 83% figure is never reported in their paper.

Moreover, several important aspects of the data are obscured by the 59%/83% calculation. First, these figures ignore the distinction that Dubowitz and colleagues drew between full disclosure and partial disclosure. Dubowitz et al. (1992) reported that partial disclosures included "suggestive doll play or an inconclusive account of alleged abuse" (p. 690). As London and colleagues noted (2008, p. 42), disclosure rates will be exaggerated if children are included in the "substantiated" group "because they demonstrate ambiguous behaviours that are considered diagnostic of abuse (e.g., suggestive doll play i.e. not diagnostic)." Although there is some debate over whether children can accurately clarify verbal disclosures with dolls (in order to identify where touching occurred), doll play would never constitute a disclosure of abuse (Poole & Bruck, 2012). Hence, including "partial disclosures" exaggerates disclosure.

Second, these figures ignore the fact that the children studied by Dubowitz and colleagues reported different degrees of medical evidence supporting a diagnosis of abuse. The authors classified "hymenal scarring" as "indicative" of abuse," and hymenal abnormalities as "compatible" with sexual abuse. Because of variability in the appearance of a normal hymen (Berenson, Heger, & Andrews, 1991), it is best to focus on the "indicative" subsample (even here, medical evidence would be stronger if the child had evidence of an STD, as discussed below; Kellogg, 2005). London and colleagues themselves warned (2008) that one should not classify children as "substantiated" if they "display 'soft' medical findings that are also present in non-abused children" (p. 42). Examination of Table 1 reveals that, of the 28 children with the strongest medical evidence of abuse (7a), 13 gave a full disclosure (1a), a disclosure rate of 46%. Dubowitz and colleagues concluded that "A sizable proportion of sexually abused children, however, do not disclose their abuse, even to skilled interviewers" (Dubowitz et al., 1992, p. 691).[1]

In their 2005 review, London and colleagues acknowledged the utility of examining disclosure rates among children with medical evidence as a means out of the circularity of relying on substantiated cases in order to estimate true disclosure rates. "[T]he conclusion that abused children do disclose abuse during formal interviews may be circular. However, there is some evidence that shows that when children are classified as abused on the basis of medical evidence or other non-child factors (confession, material evidence), most of these children do disclose abuse" (p. 217). They then proceeded to claim, erroneously, that "In Dubowitz et al. (1992), the finding that 83% of

TABLE 1 Likelihood of abuse, disclosure, and medical evidence from Dubowitz et al. (1992)

	Medical evidence			
	A. Indicative of abuse	B. Compatible with abuse	C. Normal exam	D. Total
1. Moderate to high: Full disclosure	13	0	18	[31]
2. Moderate to high: Partial disclosure	8	5	8	[21]
3. Moderate to high: No disclosure	7	3	1	[11]
4. Low to possible: Full disclosure	0	0	1	[1]
5. Low to possible: Partial disclosure	0	0	6	[6]
6. Low to possible: No disclosure	0	1	28	[29]
7. Total	[28]	—	—	99

children disclosed abuse was based on the calculation of the number of children with medical findings." (p. 217). An accurate assessment of the data reveals that over half of the children with medical evidence of abuse failed to disclose abuse.

4 | ELLIOTT AND BRIERE (1994)

The purpose of this study was to assess the "role of denial, maternal support, and symptomatology in forensic evaluations" (Elliott & Briere, 1994, p. 261). The study included 399 eight- to 15-year-old children evaluated for suspected sexual abuse. The team classified children into groups designated as abused, non-abused, or "unclear." Children's reports were coded for whether they disclosed (as opposed to failed to disclose, denied, or recanted) and whether their reports were credible, partially credible, or non-credible. The authors noted that the "evaluation relies heavily on the minor's statements… as well as on other forms of evidence when available" (pp. 261-262). The authors also reported whether there was external evidence of abuse, which included "diagnostic medical, confessions, eyewitness to abuse [or] other evidence (e.g. pornographic pictures, description of room child should not have seen)" (pp. 263-264).

Table 2 contains numbers derived from the text of the article (Elliott & Briere, 1994, p. 265). As with Table 1, numbers derived from the text are in brackets. London et al. (2005, 2008) calculated an overall disclosure rate of 57% ((1c + 2c + 4c)/9c). Inexplicably, they included recantations; the correct figure is 52% ((1c + 2c)/9c). They calculated a "substantiated" disclosure rate of 85% ((1c + 2c)/5c). To be precise, the correct figure is 84%, but this reflects a rounding error. We will use the correct figures (52%/84%), but the same points could be made with London and colleagues' calculations (57%/85%).

As with the study by Dubowitz and colleagues, the numbers make it appear that the greater the certainty of abuse, the higher the disclosure rate. However, it again illustrates the dependence of substantiation on disclosure. Children who disclosed but lacked external evidence were usually included in the abused group (1b and 2b), whereas non-disclosing and recanting children who lacked external evidence were never included (3b and 4b). Hence, disclosure increased the likelihood of substantiation, and substantiated cases therefore had higher rates of disclosure.

As with Dubowitz and colleagues, Elliott and Briere (1994) cannot be accused of circular logic. They neither reported the 52%/84% figures, nor would they interpret these numbers as estimates of the likelihood that abused children disclose. Reporting an 84% disclosure rate for substantiated abuse, as London and colleagues did, obscures the original authors' approach. First, children were interviewed up to three times as part of the assessment (Elliott & Briere, 1994, p. 264). Hence a child could initially deny abuse, but subsequently disclose, and this would be counted as a disclosure. Second, London and colleagues collapsed "credible" and "partially credible" reports, which increases the overall disclosure rate from 37% (1c/9c) to 52% and the substantiated disclosure rate from 60%

TABLE 2 Likelihood of abuse, disclosure, and external evidence from Elliott and Briere (1994)

	Evidence type		
	A. External evidence	B. No external	C. Total
1. Abused: Disclose credible	51	[99]	149
2. Abused: Disclose partially credible	28	[32]	60
3. Abused: Non-disclosing	19	[0]	19
4. Abused: Recanter	20	[0]	20
5. Abused: Total (1+2+3+4)	118	[131]	248
6. Non-abused: Non-credible recanter	—	2	2
7. Non-abused credible denial	—	70	70
8. Unclear: Disclose non-credible or denial non-credible	—	79	79
9. Grand Total (5+6+7+8)	118	[282]	399

(1c/5c) to 84%. As Elliott and Briere (1994) explained, "Disclosures were considered partially credible if a contextually embedded account of abuse was reported combined with any of the following: (a) fewer details than expected were provided; (b) the child was highly avoidant of certain topics (e.g. refused to answer or responded with 'I don't know' to more than a few questions); or (c) there was external evidence that contradicted some aspect of the child's statements, but did not negate the veracity of the reported abuse" (p. 264).

London et al. (2005) criticized the study's conclusions, arguing that "even when researchers themselves find low rates of denials or recantations, they still maintain that these are consistent with the popular view… 'that disclosing sexual abuse is more an ongoing process than a single event'" (p. 219, quoting Elliott & Briere, 1994, p. 274). However, read in context, Elliott and Briere (1994) were commenting on the high rate of "partially credible" reports: "Children who made less than completely credible statements were frequently those who initially disclosed fondling, but for whom there was evidence of more severe abuse (e.g., penetration). When provided with this evidence in a second interview, these children often made more complete statements" (p. 274).

As noted above, London et al. (2005) recognized that the circular logic of relying on cases substantiated through disclosure could be partially overcome by examining cases with non-disclosure corroborative evidence of abuse. The numbers of Elliott and Briere (1994) allow one to calculate disclosure rates among children with "external" evidence of abuse. Although the evidence in these cases was not entirely independent of disclosure (e.g., descriptions of the defendant's room, confessions), the ground truth and substantiation bias problems are reduced to some extent. Among cases with external evidence, the rates were 43% credible disclosures (1a/5a), 24% partially credible disclosures (2a/5a), 16% no disclosure (3a/5a), and 17% recantation (4a/5a). In contrast, London et al. (2005) reported that "there were 118 children involved in cases with external evidence: 84% of these 118 children *at one point* disclosed abuse" (p. 217) (emphasis added), thus including recantations in their calculation of disclosure ((1a + 2a + 4a)/5a). An accurate assessment of the data reveals that about a third (33%) of the children with external evidence of abuse denied or recanted abuse ((3a + 4a)/5a).

5 | KEARY AND FITZPATRICK (1994)

The purpose of the study was to compare disclosure rates among children who disclosed prior to assessment and children who had not previously disclosed. The study included 251 children assessed for sexual abuse, most of whom were 10 years of age or under. As the authors hypothesized, whether children disclosed was highly dependent on whether they had previously disclosed.

The numbers in Table 3 are drawn from a table showing the relation between disclosure and the outcome of the assessment. London et al. (2005, 2008) calculated an overall disclosure rate of 50% ((Rows 1 + 3)/Row 5) and a substantiated disclosure rate of 91% (1/(1+2)). Clearly, substantiation is dependent upon disclosure.

Keary and Fitzpatrick (1994) did not interpret their data in a circular fashion. They did not report the 50% figure, and did not interpret the 91% figure as a true disclosure rate. As they concluded, "Disclosure of sexual abuse during investigation was strongly positively correlated with abuse being regarded as confirmed" (p. 543). Indeed, of the disclosures, 94% were confirmed (1/(1 + 3)), and of the non-disclosures only 10% were confirmed (2/(2 + 4)).

The 50%/91% figures obscure other evidence of denial and reluctance. The assessment included up to three interviews (p. 544); the rate at which children initially denied but ultimately disclosed abuse during a subsequent interview was not reported. Because Keary and Fitzpatrick report the rate at which children disclosed abuse prior to evaluation, another flaw in only examining substantiated cases is revealed. We will refer to this as the "prior disclosure" problem. The figures ignore children who disclosed prior to assessment but recanted at the assessment. Keary and Fitzpatrick report this was true among 14% of the sample, and the percentage was particularly high among the children under 5 (41%). Because non-disclosure prevented the researchers from substantiating abuse (there is no mention in the study of other evidence of abuse), they were unable to determine whether these cases constituted recantation of a true allegation, recantation of a false allegation, or misinterpretation of the child's pre-assessment reports.

London et al. (2005) asserted that "When children have made a prior allegation but do not repeat it during a formal investigation, this should not be categorized as a recantation because it is possible that the child's first allegation was incorrect or misinterpreted, and the report during the formal investigation is accurate" (London et al., 2005, p. 209, n. 5). Hence, they chose to assume, based on this "possibility," that all pre-investigation disclosures (in this and other studies) were either false or misinterpreted.

6 | GORDON AND JAUDES (1996)

The purpose of the study was to evaluate the consistency of disclosure when children were first questioned in the emergency room regarding suspected sexual abuse and when they were subsequently evaluated by a multidisciplinary team. Specifically, the authors focused on whether the child disclosed the identity of the perpetrator across interviews. The sample included 141 three- to 14-year-old children assessed for sexual abuse.

The numbers in Table 4 are calculated based on a table from the paper identifying the characteristics of the children falling into different patterns of disclosure across the emergency room interview and the multidisciplinary team interview (Gordon & Jaudes, 1996, Table 2, p. 319). The numbers in brackets had to be calculated using the percentages and group totals shown in parentheses. London and colleagues, treating team interview disclosures as the disclosure, calculated an overall disclosure rate of 73% ((1a + 3a + 6a + 8a)/10a). They treated cases "indicated by the state" as the substantiated rate, and calculated a rate of 77% ((1a + 3a)/5a) (The precise rate is 76%.)

Gordon and Jaudes (1996) reported neither the 73% nor the 76% rate and did not treat the 76% rate as a true rate of disclosure. Instead, they began their discussion by emphasizing that "The history obtained from the child is the key component in the evaluation of suspected sexual abuse…. The information elicited and documented during history-gathering is often the only evidence obtained" (p. 320). With respect to substantiation of abuse by the state, they concluded that "[t]he ability of the state to conclude officially that sexual abuse had occurred was much higher when the child identified the alleged perpetrator" in the team interview (p. 319). Hence, substantiation was highly dependent on disclosure.

The lack of a large jump in disclosure rates between the overall disclosure rate and the substantiated disclosure rate is due to the selection of the sample. In order to be included, children had to have been evaluated as "probably victims of sexual abuse" given the results of the emergency room interview. Indeed, the state substantiated 84% of the cases (5a/10a).

TABLE 3 Likelihood of abuse and disclosure from Keary and Fitzpatrick (1994)

1. Confirmed disclosure	117
2. Confirmed non-disclosure	12
3. Unconfirmed disclosure	7
4. Unconfirmed non-disclosure	114
5. Total	250

TABLE 4 Gordon and Jaudes's (1996) state substantiation, disclosure across two interviews, and sexually transmitted disease (STD)

	A. All subjects	B. STD
1. Indicated by state: Disclosed both interviews	[78] (0.91 × 86)	[3] (0.04 × 86)
2. Indicated by state: Recanted	[5] (0.29 × 17)	[3] (0.18 × 17)
3. Indicated by state: Initial denial	[12] (0.71 × 17)	[3] (0.18 × 17)
4. Indicated by state: Never disclosed	[13] (0.62 × 21)	[5] (0.24 × 21)
5. Indicated by state: Total (1+2+3+4)	[118]	–
6. Not indicated by state: Disclosed both interviews	[8]	–
7. Not indicated by state: Recanted	[12]	–
8. Not indicated by state: Initial denial	[5]	–
9. Not indicated by state: Never disclosed	[8]	–
10. Grand Total (5+6+7+8+9)	141	14

The focus by London and colleagues on substantiated cases ignores the prior disclosure problem. A notable finding in the study is the number of children excluded from the figures of London and colleagues because they disclosed, either to their parents or the emergency room doctor, but then recanted when questioned by the team. Table 1 of Gordon and Jaudes reveals that 16% (17/103) of children with an initial disclosure failed to disclose again. Again, these rates were dismissed by London and colleagues, who stated that "[w]e do not report Gordon and Jaudes's (1996) 'recantation' rate because the child was not interviewed under the same clinical watch, but rather the first interview was a brief medical screening. Also, the authors include parents' disclosures (i.e., as historian) in the base rate" (London et al., 2005, Table 2, p. 206; cf. London et al., 2008, Table 2). As with their interpretation of the work of Keary and Fitzpatrick (1994), they assumed that all parents' reports that the child had disclosed were either misinterpretations or false, and going further ignored disclosures to the emergency room doctor.

Gordon and Jaudes (1996) were able to examine the subsample of cases with external evidence of abuse. Only 43% of the 14 children with a sexually transmitted disease disclosed at the multidisciplinary team interview ((1b + 3b)/10b). An accurate assessment of the data thus shows that, of the children with clearest evidence of abuse, most initially denied, recanted, or never disclosed.

7 | DIPIETRO, RUNYAN, AND FREDRICKSON (1997)

The purpose of the study was to identify factors predicting children's disclosure when examined for suspected sexual abuse in a hospital clinic. The authors examined 179 children one to 22 years of age ($M = 7.5$). The authors noted that the overall disclosure rate was 47%, and found higher disclosure rates among children who had

previously disclosed and who were older. London and colleagues report the 47% overall disclosure rate, and report the rate among substantiated cases as 76%. As they explain, "DiPietro et al. (1997) classified each of the children in their sample who were assessed because of suspicions of CSA as unfounded, possible, probable, or definitive abuse. Rates of disclosure during the first visit increased as a function of abuse certainty, with 7%, 8%, 59%, and 76%, respectively, disclosing" (London et al., 2005, p. 211).

DiPietro et al. did not reason about their data in a circular fashion. Rather they emphasized in the introduction that "physical signs of abuse are often inconclusive, so the identification of sexually abused children frequently relies on the disclosure of abuse by the victimized child during the medical history" (p. 134) and reiterated in the discussion that "the medical opinion of certainty of abuse was related to disclosure which would be expected in that a clear history is a major contributor to diagnoses" (p. 140). Without other evidence, "we cannot accurately assess whether the lack of a disclosure represents a false negative" (p. 141).

The data of DiPietro et al. also address the prior disclosure problem. The substantiated disclosure figures calculated by London and colleagues obscure the inconsistencies in disclosure occurring before the evaluation. Of the children who had disclosed abuse prior to being seen by the clinic, 28% failed to disclose in the clinic interview (p. 138).

8 | DEVOE AND FALLER (1999)

The purpose of the study was to examine the factors that influence whether and to what extent children evaluated for sexual abuse disclose abuse. The researchers examined 76 five- to 10-year-olds ($M = 6.8$). They concluded that "[w]ith one exception, children in this study did not disclose spontaneously about alleged sexual abuse" (p. 224).

In contrast, London and colleagues calculated that the findings in this study indicated a 62% overall disclosure rate and an 87% substantiated disclosure rate (London et al., 2005, Table 2, pp. 206, 211; London et al., 2008, p. 42). The 62% rate is the proportion of children who "described sexual activity" in the first of two interviews (DeVoe & Faller, p. 223). The 87% rate is the percentage of children "substantiated by the clinic" who disclosed in the first interview (DeVoe & Faller, p. 224).

London et al. (2008) claimed that the 47 substantiated cases were those "with corroborative evidence (e.g., medical findings, material evidence, offender confession, and offender conviction)" (p. 42). In fact, the study indicates (DeVoe & Faller, 1999, Table 3, p. 222) that corroborative evidence of that sort occurred in only 25 cases. Instead, the corroborative evidence discussed with respect to the substantiated cases was "corroborative evidence from previous reports" (Devoe & Faller, 1999, p. 224), specifically "reports of child statements in other contexts" (p. 221). This evidence is obviously dependent on disclosure. Indeed, after noting the 87% figure, Devoe and Faller (1999) emphasized that "only one [of the six substantiated cases who failed to disclose at the first interview] had not disclosed at some point during the evaluation," and that child had previously disclosed on video. Hence, a clear disclosure was essential for substantiation.

Unfortunately, Devoe and Faller (1999) did not report the disclosure rate among the 14 children with medical evidence, nor did they identify the quality of the medical evidence. Instead the external evidence included as London and colleagues noted confessions and convictions, and as London and colleagues failed to mention "police substantiation" (Devoe & Faller, 1999, Table 3, p. 222). These types of external evidence are likely highly dependent on disclosure.

Once more, the substantiated disclosure figures obscure the prior disclosure problem. Devoe and Faller (1999) report that 30% of children having made a prior disclosure failed to disclose during the first interview (17/56); 18% persisted in denying abuse at the second interview (10/56) (p. 222). As DeVoe and Faller noted, these cases could not be substantiated unless the child had disclosed on video at a previous time.

9 | SUMMARY

To summarize the argument thus far, a close reading of the six studies demonstrates how calculating disclosure rates based on substantiated cases inflates disclosure rates because substantiation is dependent upon disclosure. In each study, the substantiated disclosure rate was indeed higher than the overall disclosure rate, as London and colleagues argued, but this was attributable to substantiation bias. Ironically, the dependence of substantiation on disclosure also undermines the argument that focusing on substantiation addresses the ground truth problem. If disclosure is the only evidence of abuse, and abuse is substantiated based on disclosure, then one cannot be sure that substantiated cases are true unless one is willing to assume that all disclosures are true.

Fortunately, several of the studies were able to examine cases with more or less independent evidence of abuse that both helped to reduce the ground truth and substantiation bias problem. They consistently found lower rates of disclosure and higher rates of denial. However, the medical evidence was not always as compelling as it could be, and suspicion bias problems remain. In the next section, we will identify strategies for overcoming the methodological problems of using child disclosure studies to understand abused children's reluctance to disclose abuse.

10 | REDUCING METHODOLOGICAL PROBLEMS BY EXAMINING CHILDREN WHOSE SEXUAL ABUSE WAS SUSPECTED AND SUBSTANTIATED WITHOUT RELYING ON DISCLOSURE

Lawson and Chaffin (1992) examined the rate of nondisclosure among children with sexually transmitted diseases (STDs), a large proportion of whom were diagnosed with gonorrhea. The authors excluded children who were so young they might have acquired the STD congenitally, thus minimizing the ground truth problem. They excluded children too young to provide a verbal disclosure of abuse, and children who were old enough to have conceivably acquired the STD through consensual sex with peers. Because gonorrhea is strong evidence of sexual abuse, substantiation bias was also minimized (Kellogg, 2005) (there may have been some dependence between the STD finding and disclosure to the extent that interviewers, aware of children's diagnosis, pressed harder for a disclosure.)

In order to minimize suspicion bias, they also excluded children for whom the presenting complaint was sexual abuse. Indeed, they excluded children for whom there were any previous suspicions of abuse. This meant that children who had been questioned about abuse and denied it were excluded, contrary to the claim by London and colleagues (2008) that the sample was comprised of "hard-core deniers" (p. 40). It also meant that children who had previously disclosed abuse were excluded.

The reader will recall that suspicion bias inflates rates of disclosure because suspicion is often triggered by disclosure. Indeed, in the six studies discussed above, the samples comprised children suspected of being sexually abused, and prior disclosure rates, when they were reported, were uniformly high (Devoe & Faller, 1999, 74%, p. 222; DiPietro et al., 1997, 62%, p. 138; Keary & Fitzpatrick, 1994, 49%, p. 545). Suspicion bias will inflate the likelihood the rate of disclosure among children medically evaluated for abuse as well, because disclosures often trigger medical evaluations. For example, in the study by Heger, Ticson, Velasquez, & Bernier (2002) of medical findings in children evaluated for sexual abuse, of the sample of children with diagnostic medical evidence of sexual abuse 82% were referred for medical evaluation only after disclosing abuse.

Returning to Lawson and Chaffin (1992), the authors found that 43% (12/28) of the children made allegations of sexual abuse during the initial interview. These results are consistent with other medical research in which abuse is initially suspected and verified without reliance on the child's disclosure. Muram, Speck and Gold (1991) examined girls for whom suspicions of abuse had not arisen but who were siblings or associates of girls known to have been abused. Of the 35 girls with medical findings specific to abuse (such as hymenal tears), 51% (18/35)

disclosed abuse when questioned. The reader will recall that Dubowitz et al. (1992) and Gordon and Jaudes (1996) also had subsamples of children with medical evidence of abuse, and although they could not control for suspicion bias also found low rates of disclosure. Indeed, discussing the low rates of disclosure among children with STDs, Gordon and Jaudes concluded "[t]his is consistent with observations by Lawson and Chaffin (1992)" (p. 320).

Review of the research on gonorrhea in children reveals that false denials are quite common, and that medical researchers have understood reluctance and denial of abuse by children for nearly a century (Pollack, 1909; see also Beilin, 1931; Cohn, Steer, & Adler, 1940; Rice, Cohn, Steer, & Adler, 1941). Lyon (2007) identified 21 studies (from 1965 to 1993) examining gonorrhea in children from which one can calculate upper bounds of abuse disclosure. Although some studies explicitly refer to disclosures by children, some refer to a "history" of abuse, which could come from a child or an adult, or a "conclusion" that abuse was involved, which may or may not involve disclosure. For example, Ingram, White, Durfee, and Pearson (1982) accepted as proof of sexual contact a "history of males isolating themselves with the children under unusual circumstances that the family believed resulted in sexual contact" (p. 995). Despite the likelihood that these factors exaggerate disclosure rates, the average rate of "disclosure" was only 43% (250/579).

Whenever possible Lyon removed children younger than three years of age, who may be too young to disclose abuse. If one excludes the studies in which it was impossible to separately analyze children three years and older, the rate of "disclosure" across the remaining studies was 42% (185/437). One might object that teenagers could have contracted gonorrhea from a peer following consensual sex (London et al., 2008). If one also excludes the studies in which it was impossible to exclude teenagers, the rate of "disclosure" among the remaining studies was 53% (126/239). However one interprets the studies, the disclosure rates persistently hover closer to 50% than 100%, evincing a substantial amount of denial.

The medical research also provides evidence of incremental disclosure. In Ingram et al. (1982), more than 5 of the 29 girls with gonorrhea named a sexual contact during the first interview, whereas 13 had done so after "further interviews" (p. 995). Hence, of those who ultimately disclosed, at least 62% (8/13) did not do so initially. (It should be noted that the study selected children on the basis of suspicions of sexual abuse, so they did not control for suspicion bias.) In Farrell, Billmire, Shamroy, and Hammond (1981), 24 children ultimately provided a history of sexual contact. At most, seven children disclosed sexual contact when seen in the emergency room. Hence, of those children who ultimately disclosed, at least 71% (17/24) failed to disclose abuse when first questioned. As the authors conclude, "Our data support that of other investigators that a history of exposure to gonorrhea is infrequently obtained during the initial interview. We have demonstrated that when these same children are hospitalized and interviewed repeatedly by a skilled and sympathetic social worker, they often do give a history of exposure" (p. 152). In Shapiro, Schubert, and Myers (1993), the authors noted that of the 10 children whose chief complaint was vaginitis at the initial visit but who ultimately disclosed abuse only one of the 10 disclosed at the initial emergency room visit (p. 343). Of course, in all these cases one worries about suggestive interviewing and the possibility that the methods used by the interviewers would lead to false allegations. But CSAAS testimony is not offered to justify coercive interviewing, but to help jurors understand why abused children might deny abuse when directly asked.

Only a small percentage of sexually abused children have STDs or other medical findings. As a result, they are likely unrepresentative of abused children in various ways (Lawson & Chaffin, 1992; London et al., 2008). This point was emphasized by Bruck in her expert report, and by the court in *J.L.G.* The critical question, however, is whether they are unrepresentative in ways that would distort children's reluctance to disclose abuse. For example, London et al. (2008) argue that gonorrhea is more common among children who experienced repeated penetrative abuse, but they offer no evidence that this would lead to the dramatically lower disclosure rates observed in these studies. Their primary argument against using STDs as substantiation is that they occur only among a small percentage of abused children, but infrequent is not the same as non-representative. The question is whether abused children with STDs are systematically less likely to disclose abuse when questioned than abused children without STDs.

With respect to representativeness of different types of study examining children's disclosure, London et al. (2008) made an unusual claim. On the one hand, they acknowledged that "by definition, because a significant

proportion of sexually abused children are not officially interviewed, the clinic samples [for example, the six "methodologically superior" studies] will not be representative of the whole population." (p. 43). As they put it, excluded are the "silent children" (p. 43). These are the children who are particularly likely to deny abuse when questioned. This implicitly acknowledges the suspicion bias problem.

On the other hand, London and colleagues argued that the clinic studies are "representative of children who are involved in investigations of CSA" (p. 43), and thus appropriate for assessment of the likelihood of denial. This argument is flawed in two respects. First, it overlooks how focusing on the substantiated subsamples in these studies increases disclosure substantiation bias. Second, it fundamentally misunderstands the argument that experts make when explaining denial to jurors. It may be the case that most abused children who show up in clinics are willing, even eager, to disclose abuse. This helps to explain why many abuse disclosures can be elicited with open-ended questions (Lamb, Brown, Hershkowitz, Orbach, & Esplin, 2018). But this says nothing about whether a case in which the child denied abuse, either inside or outside the clinic, should be viewed with skepticism. In order to answer that question one needs to assess the reluctance of abused children to disclose when selection biases are eliminated.

11 | RECENT EVIDENCE OF ABUSE DENIAL

More recent research has examined rates of disclosure, and supports the argument made here: as problems of ground truth, suspicion bias and substantiation bias are reduced, false denials of abuse remain a major problem. Hershkowitz et al. (2014) examined over 400 children referred for suspicions of physical or sexual abuse. The study reduced ground truth and disclosure substantiation bias problems by limiting the sample to cases with corroborative evidence such as "suspects' admissions, disinterested eyewitness testimony, medical evidence (including observable physical injuries), and material evidence" (p. 339). The study utilized the NICHD structured protocol (indeed, most of the children received a revised version of the protocol, with added encouragements designed to increase disclosure), a state of the art approach to non-suggestive interviewing (Lamb et al., 2018). The study thus addresses claims that denials are only a problem when children are poorly interviewed. However, the study could not control for suspicion bias (and 34% of children had previously disclosed). Hence, to the extent that children's disclosures led to the investigation, disclosure rates were exaggerated.

The overall disclosure rates were 57% for physical abuse and 44% for sexual abuse. Hence, about half of children denied abuse. The study was also able to address the prior disclosure problem, because if children had made a prior disclosure and failed to disclose during the assessment this could be considered recantation of a true case of abuse. Among children who had previously disclosed, the disclosure rate was 71%. This means that 29% of children recanted their abuse.

Malloy and colleagues (Malloy, Lyon, & Quas, 2007; Malloy, Mugno, Rivard, Lyon, & Quas, 2016) examined recantation in 257 dependency cases of sexual abuse substantiated by social services investigation. The authors had information about case characteristics and could address the ground truth problem. The study could not control for suspicion bias or substantiation bias. If children's disclosures led to investigation, or children's denials shut down investigations, then more reluctant children were excluded, and recantation rates underestimated. Contrary to the supposition that higher rates of recantation suggest that allegations are false, Malloy and colleagues found that recantation rates were no higher among cases involving possible custody disputes. Similarly, contrary to claims that recantation rates are lower when abuse is more certain, recantation rates were not lower among cases with corroborative evidence. The researchers found a 23% rate of recantation, and recantations were more common among younger children, children abused by a parent figure, and children lacking familial support. These findings are consistent with expert testimony explaining how abused children are susceptible to pressures to recant.

Another promising approach is to examine cases for which there is photographic or digital evidence of abuse, which can establish ground truth, reduce substantiation bias, and even solve the suspicion bias problem if the digital

evidence is discovered before the child is questioned about abuse. Studies examining documented abuse have found high rates of reluctance and denial (Katz, Piller, Glucklich, & Matty, 2018; Leander, 2010; Sjöberg & Lindblad, 2002).

12 | DISCUSSION

Close examination of the "methodologically superior" studies cited by the experts in *J.L.G.* demonstrates the fallacy of estimating abused children's willingness to disclose abuse on substantiated cases. Much of the research fails to solve the ground truth problem, because disclosures can be false, and it also falls prey to the suspicion and substantiation bias problems, because abuse is often suspected because of a disclosure, and because substantiation is highly dependent on disclosure.

If instead one identifies cases with independent evidence of abuse, that is, evidence that does not result from or depend on disclosure, one can address the ground truth, suspicion bias, and substantiation bias problems. When this is done, one finds what practitioners have understood for over a century: the fact that a child has at one point denied or recanted abuse should not lead one to assume the allegation is false. Instead, denials and recantations are understandable reactions to initial suspicions and subsequent questioning about abuse.

At the same time, suspicion bias means that children seen for evaluation of sexual abuse are likely to be uncommonly willing to disclose abuse. Researchers have demonstrated that, as a result, open-ended questions are highly effective in eliciting disclosures from most children who have previously disclosed (Lamb et al., 2018). However, when children do not readily disclose, professionals should not be too quick to dismiss suspicions of abuse as unfounded. An active area of research is exploring supportive techniques that can overcome children's reluctance to disclose abuse without increasing false allegations (Blasbalg, Hershkowitz, & Karni-Visel, 2018).

The long-term effects of *J.L.G.* on CSAAS testimony are yet to be determined. The New Jersey Supreme Court has applied "pipeline retroactivity" to the holding, which means that cases with appeals pending have taken advantage of the ban on expert testimony (*New Jersey v. J.E.P.*, 2020), and a number of convictions have been reversed (e.g., *New Jersey v. J.J.R.*, 2020; *New Jersey v. U.M.*, 2020).

The case has been cited by lower courts in other states (*California v. Harrison*, 2020; *California v. Munch*, 2020; *Missouri v. Marshall*, 2020), but thus far the courts have rejected the arguments and upheld the admissibility of CSAAS testimony. Probably because *J.L.G.* made no mention of how methodological problems conceal reluctance and denial, none of the issues documented here have been discussed. In order for future courts to accurately assess the scientific status of CSAAS, it is important for researchers to apprise them of these methodological concerns.

ACKNOWLEDGMENTS

Preparation of this article was supported in part by NICHD Grant HD087695. In the evidentiary hearing in *J.L.G.*, the first author testified pro bono for the prosecution as an expert witness.

ENDNOTE

[1] Unlike the other five studies described here, Dubowitz et al. (1992) was specifically discussed by the court in *J.L.G.* Unfortunately, its treatment of the study only highlighted difficulties in interpreting numbers. In the portion of the opinion finding scientific support for the proposition that children delay reporting, the court noted that only 21% of the sample who were judged to be moderately or highly likely to have been abused both fully disclosed and had medical findings indicative of abuse (1a/(1d+2d+3d) or 13/63). The low percentage, however, says as much about the prevalence of medical findings as it does about children's reluctance to disclose, and of course it speaks to denial rather than delay.

REFERENCES

Alaggia, R., Collin-Vézina, D., & Lateef, R. (2019). Facilitators and barriers to child sexual abuse (CSA) disclosures: A research update (2000–2016). *Trauma, Violence, & Abuse*, *20*, 260–283.

Azzopardi, C., Eirich, R., Rash, C. L., MacDonald, S., & Madigan, S. (2019). A meta-analysis of the prevalence of child sexual abuse disclosure in forensic settings. *Child Abuse & Neglect*, *93*, 291–304. https://doi.org/10.1016/j.chiabu.2018.11.020

Beilin, L. M. (1931). Gonorrheal urethritis in male children (with some observations on their sexual impulses). *Journal of Urology*, *25*, 69–84. https://doi.org/10.1016/S0022-5347(17)72824-8

Berenson, A., Heger, A., & Andrews, S. (1991). Appearance of the hymen in newborns. *Pediatrics*, *87*, 458–465. Retrieved from https://pediatrics.aappublications.org/content/87/4/458

Blasbalg, U., Hershkowitz, I., & Karni-Visel, Y. (2018). Support, reluctance, and production in child abuse investigative interviews. *Psychology, Public Policy, and Law*, *24*, 518–527.

Bruck, M. (2017), Psychological Report, State of New Jersey v. J.L.G.

Bruck, M., & Ceci, S. J. (2004). Forensic developmental psychology: Unveiling four common misconceptions. *Current Directions in Psychological Science*, *13*, 229–232. https://doi.org/10.1111/j.0963-7214.2004.00314.x

Bruck, M., & Ceci, S. J. (2009). Reliability of child witnesses' reports. In J. L. Skeem, K. S. Douglas, & S. O. Lilienfeld (Eds.), *Psychological science in the courtroom: Consensus and controversy* (pp. 149–171). New York, NY: Guilford Press.

Bruck, M., Ceci, S. J., & Principe, G. F. (2007). The child and the law. In K. A. Renninger (Ed.), *Handbook of child psychology, Vol. IV. Child psychology in practice* (pp. 776–816). Hoboken, NJ: Wiley.

California v. Harrison (2020). WL 5105112 (Cal. App. 2020).

California v. Munch (2020). 266 Cal.Rptr.3d 136 (Ca. App. 2020).

Ceci, S. J., & Bruck, M. (2013). Loftus's Lineage in developmental forensic research: Six scientific misconceptions about children's suggestibility. In M. Garry & H. Hayne (Eds.), *Do justice and let the sky fall: Elizabeth F. Loftus and her contributions to science, law, and academic freedom* (p. 65). Mahwah, NJ: Psychology Press.

Ceci, S. J., Kulkofsky, S., Klemfuss, J. Z., Sweeney, C. D., & Bruck, M. (2007). Unwarranted assumptions about children's testimonial accuracy. *Annual Review of Clinical Psychology*, *3*, 311–328. https://doi.org/10.1146/annurev.clinpsy.3.022806.091354

Cohn, A., Steer, A., & Adler, E. L. (1940). Gonococcal vaginitis: A preliminary report on one year's work. *Venereal Disease Information*, *21*, 208–220.

Connecticut v. Favoccia. (2012). 51 A.3d 1002 (Conn. 2012).

DeVoe, E. R., & Faller, K. C. (1999). The characteristics of disclosure among children who may have been sexually abused. *Child Maltreatment*, *4*, 217–227. https://doi.org/10.1177/1077559599004003003

DiPietro, E. K., Runyan, D. K., & Fredrickson, D. D. (1997). Predictors of disclosure during medical evaluation for suspected sexual abuse. *Journal of Child Sexual Abuse*, *6*, 133–142. https://doi.org/10.1300/J070v06n01_09

Dubowitz, H., Black, M., & Harrington, D. (1992). The diagnosis of child sexual abuse. *American Journal of Diseases of Children*, *146*, 668–693. https://doi.org/10.1001/archpedi.1992.02160180046015

Elliott, D. M., & Briere, J. (1994). Forensic sexual abuse evaluations of older children: Disclosures and symptomatology. *Behavioral Sciences & the Law*, *12*, 261–277. https://doi.org/10.1002/bsl.2370120306

Faller, K. C., & Henry, J. (2000). Child sexual abuse: A case study in community collaboration. *Child Abuse & Neglect*, *24*, 1215–1225. https://doi.org/10.1016/S0145-2134(00)00171-X

Farrell, M. K., Billmire, E., Shamroy, J. A., & Hammond, J. G. (1981). Prepubertal gonorrhea: A multidisciplinary approach. *Pediatrics*, *67*, 151–153.

Finkelhor, D. (1979). *Sexually victimized children*. New York, NY: Free Press.

Finkelhor, D. (1980). Risk factors in the sexual victimization of children. *Child Abuse & Neglect*, *4*, 265–273

Gagnon, J. H. (1965). Female child victims of sex offenses. *Social Problems*, *13*, 176–192.

Gordon, S., & Jaudes, P. K. (1996). Sexual abuse evaluation in the emergency department: Is the history reliable? *Child Abuse & Neglect*, *20*, 315–322. https://doi.org/10.1016/0145-2134(96)88716-3

Gray, E. (1993). *Unequal justice: The prosecution of child sexual abuse*. New York, NY: Free Press.

Groth, A. N. (1978). Patterns of sexual assault against children and adolescents. In A. Burgess, A. N. Groth, L. Holmstrom, & S. Sgroi (Eds.), *Sexual assault of children and adolescents* (pp. 3–24). Lexington, MA: Lexington.

Heger, A., Ticson, L., Velasquez, O., & Bernier, R. (2002). Children referred for possible sexual abuse: Medical findings in 2384 children. *Child Abuse & Neglect*, *26*, 645–659. https://doi.org/10.1016/S0145-2134(02)00339-3

Hershkowitz, I., Lamb, M. E., & Katz, C. (2014). Allegation rates in forensic child abuse investigations: Comparing the revised and standard NICHD protocols. *Psychology, Public Policy, & Law*, *20*, 336–344. https://doi.org/10.1037/a0037391

Ingram, D. L., White, S. T., Durfee, M. F., & Pearson, A. W. (1982). Sexual contact in children with gonorrhea. *American Journal of Diseases of Children*, *136*, 994–996. https://doi.org/10.1001/archpedi.1982.03970470038010

Katz, C., Piller, S., Glucklich, T., & Matty, D. E. (2018). "Stop waking the dead": Internet child sexual abuse and perspectives on its disclosure. *Journal of Interpersonal Violence*. https://doi.org/10.1177/0886260518796526

Keary, K., & Fitzpatrick, C. (1994). Children's disclosure of sexual abuse during formal investigation. *Child Abuse & Neglect*, *18*, 543–548. https://doi.org/10.1016/0145-2134(94)90080-9

Kellogg, N. (2005). The evaluation of sexual abuse in children. *Pediatrics*, *116*, 506–512. https://doi.org/10.1542/peds.2005-1336

Lamb, M. E., Brown, D. A., Hershkowitz, I., Orbach, Y., & Esplin, P. W. (2018). *Tell me what happened: Structured investigative interviews of child victims and witnesses* (2nd ed.). London, UK: Wiley.

Lawson, L., & Chaffin, M. (1992). False negatives in sexual abuse disclosure interviews: Incidence and influence of caretakers' belief in abuse in cases of accidental abuse discovery by diagnosis of STD. *Journal of Interpersonal Violence*, *7*, 532–542. https://doi.org/10.1177/088626092007004008

Leander, L. (2010). Police interviews with child sexual abuse victims: Patterns of reporting, avoidance and denial. *Child Abuse & Neglect*, *34*, 192–205.

Leclerc, B., Proulx, J., & Beauregard, E. (2009). Examining the modus operandi of sexual offenders against children and its practical implications. *Aggression and Behavior*, *14*, 5–12.

Lemaigre, C., Taylor, E. P., & Gittoes, C. (2017). Barriers and facilitators to disclosing sexual abuse in childhood and adolescence: A systematic review. *Child Abuse & Neglect*, *70*, 39–52.

London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, & the Law*, *11*, 194–226. https://doi.org/10.1037/1076-8971.11.1.194

London, K., Bruck, M., Wright, D. B., & Ceci, S. J. (2008). Review of the contemporary literature on how children report sexual abuse to others: Findings, methodological issues, and implications for forensic interviewers. *Memory*, *16*, 29–47. https://doi.org/10.1080/09658210701725732

London, K., & Kulfofsky, S. (2009). Factors affecting the reliability of children's forensic reports. In G. M. Davies & D. B. Wright (Eds.), *Current issues in applied memory research* (pp. 119–141). New York, NY: Psychology Press.

Lyon, T. D. (2007). False denials: Overcoming methodological biases in abuse disclosure research. In M. Pipe, M. Lamb, Y. Orbach, & A. Cederborg (Eds.), *Disclosing abuse: Delays, denials, retractions and incomplete accounts* (pp. 41–62). Mahwah, NJ: Erlbaum.

Lyon, T. D. (2009). Abuse disclosure: What adults can tell. In B. Bottoms, C. Najdowski, & G. Goodman (Eds.), *Children as victims, witnesses, and offenders: Psychological science and the law* (pp. 19–35). New York, NY: Guilford.

Malloy, L. C., Lyon, T. D., & Quas, J. A. (2007). Filial dependency and recantation of child sexual abuse allegations. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 162–170. https://doi.org/10.1097/01.chi.0000246067.77953.f7

Malloy, L. C., Mugno, A. P., Rivard, J. R., Lyon, T. D., & Quas, J. A. (2016). Familial influences on recantation in substantiated child sexual abuse cases. *Child Maltreatment*, *21*, 256–261.

McGuire, K., & London, K. (2020). A retrospective approach to examining child abuse disclosure. *Child Abuse & Neglect*, *99*, 104263.

Missouri v. Marshall (2020). 596 S.W.3d 156, 162 (Mo. App. 2020).

Muram, D., Speck, P. M., & Gold, S. S. (1991). Genital abnormalities in female siblings and friends of child victims of sexual abuse. *Child Abuse & Neglect*, *15*, 105–110. https://doi.org/10.1016/0145-2134(91)90095-U

New Jersey v. J.E.P. (2020). 235 A.3d 157 (NJ).

New Jersey v. J.J.R. (2020). WL 2071469 (N.J. App. 2020).

New Jersey v. J.L.G. (2018). 190 A.3d 442 (N.J.).

New Jersey v. J.Q. (1993). 617 A.2d 1196 (N.J.).

New Jersey v. J.R. (2017). 152 A.3d 180 (N.J.).

New Jersey v. R.B. (2005). 873 A.2d 511 (N.J.).

New Jersey v. U.M. (2020). WL 5079890 (N.J. App. 2020).

New York v. Spicola (2011). 947 N.E. 2d 620 N.Y.

O'Donohue, W., & Benuto, L. (2012). Problems with child sexual abuse accommodation syndrome. *Scientific Review of Mental Health Practice, 9*, 20–28.

Pollack, F. (1909). The acquired venereal diseases in children: A report of 187 children treated in the women's venereal department of the Johns Hopkins hospital dispensary. *Johns Hopkins Hospital Bulletin*, *218*, 142–149. http://dx.doi.org/10.1097/00000441-190910000-00045

Poole, D. A., & Bruck, M. (2012). Divining testimony? The impact of interviewing props on children's reports of touching. *Developmental Review*, *32*, 165–180. https://doi.org/10.1016/j.dr.2012.06.007

Rice, J. L., Cohn, A., Steer, A., & Adler, E. L. (1941). Recent investigations on gonococcic vaginitis. *Journal of the American Medical Association*, *117*, 1766–1769. https://doi.org/10.1001/jama.1941.02820470014004

Rush, E., Lyon, T. D., Ahern, E. C., & Quas, J. A. (2014). Disclosure suspicion bias and abuse disclosure: Comparisons between sexual and physical abuse. *Child Maltreatment*, *19*, 113–118. https://doi.org/10.1177/1077559514538114

Russell, D. E. H. (1983). The incidence and prevalence of intrafamilial and extrafamilial sexual abuse of female children. *Child Abuse & Neglect*, *7*, 133–146.

Shapiro, R. A., Schubert, C. J., & Myers, P. A. (1993). Vaginal discharge as an indicator of gonorrhea and chlamydia infection in girls under 12 years old. *Pediatric Emergency Care*, *9*, 341–345.

Sjöberg, R. L., & Lindblad, F. (2002). Limited disclosure of sexual abuse in children whose experiences were documented by videotape. *American Journal of Psychiatry*, *159*, 312–314.

Stroud, D. D., Martens, S. L., & Barker, J. (2000). Criminal investigation of child sexual abuse: A comparison of cases referred to the prosecutor to those not referred. *Child Abuse & Neglect*, *24*, 689–700. https://doi.org/10.1016/S0145-2134(00)00131-9

Summit, R. C. (1983). The child sexual abuse accommodation syndrome. *Child Abuse & Neglect*, *7*, 177–193.

Summit, R. C. (1992). Abuse of the child sexual abuse accommodation syndrome. *Journal of Child Sexual Abuse*, *1*, 153–163.

Williams, S., McWilliams, K., & Lyon, T. (2020). Children's concealment of a minor transgression: The role of age, maltreatment, and executive functioning. *Journal of Experimental Child Psychology*, *191*, 104664

Zajac, R., Garry, M., London, K., Goodyear-Smith, F., & Hayne, H. (2013). Misconceptions about childhood sexual abuse and child witnesses: Implications for psychological experts in the courtroom. *Memory*, *21*, 608–617. https://doi.org/10.1080/09658211.2013.778287

How to cite this article: Lyon TD, Williams S, Stolzenberg SN. Understanding expert testimony on child sexual abuse denial after *New Jersey v. J.L.G.*: Ground truth, disclosure suspicion bias, and disclosure substantiation bias. *Behav Sci Law*. 2020;38:630–647. https://doi.org/10.1002/bsl.2490