

On the Assessment of Children in Suspected Child Sexual Abuse in Light of Daubert and Frye: Limitations of Profiles and Interviews as Scientifically Grounded Evidence*

REFERENCE: Bigelow BJ. On the assessment of children in suspected child sexual abuse in light of Daubert and Frye: limitations of profiles and interviews as scientifically grounded evidence. *J Forensic Sci* 2000;45(3):573–581.

ABSTRACT: Practice with children and families entails the higher probability of encountering forensic issues of child sexual abuse (CSA) assessments for which relatively few psychologists, allied mental health and legal practitioners are sufficiently well equipped. The current paper reviews some of the key psycholegal issues bearing on the assessment of suspected CSA in the contexts of: (a) recent psycholegal precedence and common law rules of reliability and admissibility of CSA profile evidence; (b) the empirical problems with CSA syndromes; and (c) the problems with children's interviews as evidence, and suggestions for valid interviewing guidelines supporting free recall. These psycholegal issues are presented in terms of the Frye standard for expert testimony and the Federal Rules of Evidence, with recent American and Canadian case illustrations, such as *Daubert v. Merrell Dow Pharmaceuticals*, *Hadden v. State of Florida* (1997), *Bighead v. The United States of America* (1997), *Diocese of Winona v. Interstate Fire & Cas. Co.* (1994), and *R. v. Simpson* (1996).

KEYWORDS: forensic science, forensic psychiatry, forensic psychology, child sexual abuse, common law, Frye, Federal Rules of Evidence, admissibility, reliability, jury bias, Daubert, Hadden, profiles, disorders, syndromes, interviews, memory

The keynote theme organizing the 21st Meeting of the American Academy of Forensic Sciences (1999) was "Who will set the standards for the 21st century?" Central to this theme was the pressing need to rise to the "Daubert challenge." The implications of Daubert as a springboard for the admissibility of expert forensic evidence permeate all aspects of forensic science and extend to behavioral science. The current paper addresses the scientific bases and limitations of child sexual abuse (CSA) assessments.

Since there has been a substantial increase in reported CSA in recent years (1), it is not surprising that public awareness of CSA has also increased. While public awareness helps to protect children, it may unfortunately also foster injustice, and there are indications that this risk is real. For example, jurors are significantly more likely to be generically prejudiced against defendants charged with

CSA (2). Given the presumption of innocence in common law and the need to protect the rights of both those accused and the alleged victims, it is especially incumbent upon mental health practitioners and allied service providers to adhere to very high scientific standards when assessing CSA. Regrettably, however, most mental health service providers and lawyers are not specifically versed in terms of valid CSA interviewing (3). Presumably then, their knowledge of valid behavioral manifestations of CSA is also likely to be deficient. Accordingly, there is a pressing need for more user-friendly literature that integrates behavioral and legal issues.

The current paper is a distillation and integration of some of the key legal issues with selected and relevant CSA literature. Such issues include but are not limited to the rules of evidence, limitations of profiles and syndromes, competent interviewing of children, and suggestions for dealing with initial disclosures. Only the assessment of the child alone is addressed here. A comprehensive assessment of CSA requires the corroboration of data from multiple sources.

Especially in child and family practice, CSA issues may be inescapable. Not only does this area of practice bring one into frequent contact with children, issues of custody and access may elevate the risk of CSA allegations. Notwithstanding the technical difficulties involved in the classification of cases (4), false or unsubstantiated allegations of CSA in custody disputes are estimated to run as high as 50% in the United States (5–7).

Given the number of potentially conflicting roles involved in working with families, such as therapist, mediator, expert advisor, and assessor (8,9), the relatively high risk of encountering disclosures of alleged CSA, and the relatively higher rate of complaints in the area of custody and access assessments, the need for psycholegal CSA knowledge is particularly acute. Partly as an acknowledgment of this state of affairs, the Ontario College of Psychologists, along with many other State psychology boards, has recently (1998) required applicants to pass a supplemental jurisprudence exam as one of the formal criteria of registration.

Uncorroborated clinical inference and contaminated interviewing (10) have no place in preparing scientifically informed CSA evidentiary material (11). Knowing the scientific limitations of one's discipline is essential in providing a balanced and useful forensic evaluation (12). When clinical descriptions of children are improperly elevated to the status of empirically contentious syndromes, disorders or even simple profiles, numerous psycholegal errors and miscarriages of justice can result because they do not sufficiently reduce uncertainty below reasonable doubt and can then bias the trier of fact and invade the province of the jury.

¹ Full Professor, Department of Psychology, Laurentian University, Sudbury, Ontario, Canada, P3E 2C6.

* This manuscript is based on a paper presented at the 51st Annual Meeting, American Academy of Forensic Sciences, Orlando, FL, February 1999.

Received 16 Feb. 1999; and in revised form 9 June, 13 July 1999; accepted 16 July 1999.

Given the substantial disruption to all family members that frequently follows the reporting of suspected CSA (12), the potential damage to children and their families of failures to legitimately intervene, and the irreparable harm and liability that flow from unfounded professional opinion, the practitioner working in a forensic CSA role bears a very heavy burden of responsibility to conduct a competent assessment. Any serious flaws in the assessment will surely become very evident within the adversarial court system. Moreover, resourceful aggrieved parents with sufficient computing skills have ready access to the same current case law and literature involving psycholegal testimony as do lawyers and expert assessors.

Recent Precedence and Admissibility of Scientific and Behavioral Testimony

It is first necessary to consider what the legal system views as scientifically useful evidence and how this applies to behavioral testimony, whether given by psychologists, psychiatrists, or social workers. Such psycholegal considerations are quite different from clinical ones in a variety of ways which turn on the rigor of our findings (i.e., probability), their admissibility in court (i.e., relevancy) and our avoidance of biasing the jury. These principles are basic to English common law in Canada (13) and have been formalized in the United States as the Federal Rules of Evidence (14). Although there is a good deal of variation in how these federal rules of evidence are interpreted and administered in courts at various state and provincial levels, these legal tools are nevertheless broad guidelines to good forensic practice in general. When science and the law work well together, there is a relatively smooth transition between the two as to the validity of data. Thus, if one's assessment is knowingly to be used in court, then it should be prepared as a scientific exhibit rendered interpretable according to common law rules of evidence. This need not be an obscure task as data that are reliable and valid are often consistent with the spirit of good evidence (15). Scientifically grounded reports can also provide a useful educational service. For example, believability and statement consistency are important legal tools but may be unfounded bases for evaluating the accuracy of children's memory (6,16).

Since clinical behavioral science is often less than reliable and valid, in actual practice, an optimal fit between testimony and legal requirements is difficult to achieve. This is especially so with respect to the assessment of CSA. Precisely because of its probabilistic nature and considerable usage of unfounded theoretical constructs, such as profiles, disorders, and syndromes, behavioral data are often technically quite removed from a hard factual basis. Theoretical constructs are the lifeblood of good scientific discovery but without a clearly valid and reliable foundation they are not factual entities. Thus, behavioral sciences like psychology are at times referred to as *junk science* (17). However, for all their ill-grounded unfairness, such barbs do us all a service as scientific rigor is our business. Wherever common law principles apply, we must be watchful of the fit between how we practice and what stands up in court.

The courts are not without their own difficulties as to what constitutes valid and reliable evidence. Two groundbreaking cases in the U.S., *Daubert v. Merrell Dow Pharmaceuticals*, 1977 (18), henceforth referred to as *Daubert*, and *Hadden v. State of Florida*, 1997 (19), henceforth referred to as *Hadden*, are of particular relevance and have resurrected debates concerning the Kelly-Frye (i.e., Frye) test (20) as a basis for the admissibility of expert scientific evidence. *Daubert* is a federal decision and *Hadden* is a state one.

Daubert dealt with biological issues and rejected Frye in favor of the Federal Rules which stress that admissible scientific evidence is based on *sound methodology*. *Hadden* invoked the Frye test as a standard and dealt with CSA profiles. Essentially, the Frye test requires that any novel procedure or process in question is admissible if it is reliable for its own purposes and is *commonly acceptable by the relevant scientific community*. Both of the above cases (*Daubert*; *Hadden*) are precedent setting and instructive because they strike at the heart of the admissibility of expert scientific testimony and underscore the need to assure the relevant factual bases of our expert opinions.

It is important to note that neither standard is clearly better than the other. On the one hand, the Frye standard of general acceptance by one's scientific community is tougher to meet than one based purely on sound scientific methodology. The gauntlet of peer review and replication is an important safeguard of the quality of good science. A novel finding may not meet this high standard yet still be scientifically valid and comply with the Federal Rules. Moreover, general acceptance is not a guarantee that a finding is methodologically sound. On the other hand, *Daubert* is stricter than Frye in requiring evidence to make a relevant fact more probable with it than without. In practice though, a probable fact is likely, although not necessarily, as in *Daubert*, one that has been well accepted by the scientific literature.

Hadden is particularly germane to CSA assessments as it addresses the admissibility of psychological profiles as legitimate syndromes, and potentially applies to all types of cases involving psychological profile evidence. Profiles or syndromes are often proffered as nontestamentary evidence of CSA victims and their underlying facts or databases often go unchallenged. However, in *Hadden*, under Federal Rule 705, the court cross examined the expert witness in order to disclose these underlying facts. A conviction of child sexual assault was overturned because CSA profile evidence (e.g., CSA accommodation syndrome (CSAAS), 21,22) was judged as not sufficiently reliable to have met the Frye test. That is, CSAAS was judged as not reliable for its own purposes and was not generally accepted by the relevant scientific community. The court left open future potential challenges to a variety of other novel syndromes such as post-traumatic stress disorder (PTSD) and pedophile profiles.

This ruling casts a very wide net indeed, potentially challenging the admissibility of personality profiles and a host of other classifications using behavioral descriptors, such as personality subtest profiles (e.g., MMPI-2, MCMI-II) and perhaps various psychiatric disorders (e.g., DSM IV). When a profile (e.g., CSAAS) is elevated to the status of a syndrome, it also implies causative etiology. Whatever the technical implications of various imprimaturs given to profiles, the courts are having difficulty in admitting them into evidence. Indeed, in their popularization of criminal profiling as an investigative policing tool, Douglas and Olshaker (23) complained bitterly of the reluctance of the courts to accept investigative *signature* profiles of serial killers as evidence for the prosecution of such cases. Perhaps at the current state of the science, CSA profile data are better regarded as loosely descriptive investigative tools rather than factual entities.

The *Daubert* (18) ruling has sent a clarion call throughout the entire forensic science community and therefore has potential impact on behavioral science and CSA evidence. In *Daubert*, the court rejected the Frye standard in favor of the Federal Rules, thereby disposing of the requirement of general acceptance of scientific evidence by the relevant scientific community. Instead, *Daubert* stressed that findings be based on the scientific method and be

probable. In *Daubert*, a previous decision against two child petitioners, based on the alleged teratogenic effects of a manufactured drug, was vacated and remanded. It was judged that the petitioners' expert evidence was unfairly rejected because its procedures didn't meet the stricter Frye standard. The court ruled that the Frye test has been superseded by the Federal Rules of Evidence (Rule 702) which speaks to expert scientific evidence, making no specific reference to the *general acceptability* of this evidence by the relevant scientific community, but does provide the judge with a liberal interpretation of what it means to be relevant and reliable. All evidence that is relevant to the jury is admissible if it helps in the determination of the issue at hand (Rule 401). Instead of the Frye standard, expert scientific evidence in federal court, and in state courts using the Federal Rules, is relevant and reliable as long as the facts or data forming the basis of an expert's opinion are of a *type* reasonably relied upon by experts in the field (Rule 703). In *Daubert* this means *grounded on scientifically informed principles and methodology* (Rule 702). According to the court's interpretation of Rule 702 in *Daubert*, scientific grounding connotes ". . . a body of known facts or ideas inferred from facts or accepted as true on good grounds" (p. 1). In summary, the need for general acceptance of an expert opinion by one's scientific community has been supplanted in federal court by an opinion that is informed by the application of sound scientific principles and methodology.

In the preceding *Daubert* decision, using the stricter Frye standard, unpublished articles reanalyzing prior data and published studies based on animal data were initially deemed not to properly apply to the matters before the court because they were not published in peer reviewed journals and were not then commonly accepted by the relevant scientific community. But the more liberal 1993 *Daubert* decision made these data admissible and the decision is now under review. What this case means to behavioral scientists is not yet entirely clear, but does suggest that our expert testimony, as scientific, will be continually subjected to challenges as to its reliability, validity and its grounding in scientific methodology.

Daubert also implies that what is commonly accepted is not necessarily admissible unless its underlying factual basis can be demonstrated. This feature is a direct challenge to unfounded expert opinions based purely on clinical experiences or interpretive constructs, no matter how widely held, which have no sound demonstrable empirical bases. The Frye test does not address technical or specialized knowledge but the Federal Rules do (Rule 702) and stipulate that scientific means knowledge by an expert qualified by "knowledge, skill, experience, training or education" (Rule 702). Here, scientific implies not mere subjective belief or unsupported speculation. The Frye test does not make this stipulation, but in *Hadden* the court clearly required expert evidence in chief to have a clear factual basis. Both these standards (i.e., Frye and the Federal Rules) address the necessity of expert scientific testimony having reliability and the Federal Rules also stipulate the parameters of the relevance of evidence (Rule 401), which are to make the existence of a consequential fact, that is one that is pertinent to the charge, more or less probable than it would be without this evidence.

The Federal Rules also stipulate that the expert testimony must not bias the jury on ultimate issues (Rule 704b), which means not interfering with the domain of the jury to arrive at its own determinations of guilt or innocence. While experts can make statements which "embrace the ultimate issue" (Rule 704a), they cannot make statements (i.e., opinions or inferences) pertaining to the defendant's "mental state or condition" in relation to the charge because intent (i.e., *mens rea*) is a vital component of guilt to be determined

by the trier of fact (i.e., judge and/or jury) alone. It appears that factually contentious profile syndrome evidence fails on both counts. Unfounded profiles which masquerade as syndromes and are accepted as nontestamentary evidence improperly embrace Rules 401 and 704a. That is, such profile evidence pretends to link symptoms to precipitating events and then unfairly points a judicial finger at the accused and therefore unfairly implies his or her guilt, and because it has an *aura of infallibility*, unfounded profiles also risk an invasion of the province of the jury (Rule 704b).

Owing to the relatively more fragile bases of behavioral science relative to the harder disciplines (12,24), it is debatable, but nevertheless potentially defensible, that personality profiles using extensively validated tests actually meet a sufficiently high objective standard (i.e., Rule 401) to warrant their use in court. That which is both psychometrically significant and legally useful is still very much a hotly debated topic. Note, however, that psychometric profiles must have sufficient predictive validity (Rule 401) so as to materially reduce uncertainty. Does a given profile or syndrome correctly classify people beyond a reasonable doubt (95% accuracy)? Is it clear and convincing (80%)? Or is it correct on balance of probabilities (>50%)? As reviewed in the next section, CSA profiles and any of their variants (e.g., CSAAS, pedophilic profile, psychic harm) fail to attain even this least stringent standard.

A key problem with putative syndrome testimony is that it implies that all people (or at least the vast majority) who have given constellations of symptoms have identifiable disorders with specific underlying etiologies (24). Unscientifically validated syndromes imply coherences and etiologies that simply are not present and can then act as metaphorical *psycho viruses* in the court's decision-making deliberations. As Rotgers and Barrett (15) so clearly argued, we cannot engage in reverse logic in which we erroneously claim that all people who suffer given types of trauma show specific behaviors unique to the trauma. This is precisely the logic that syndrome evidence indulges in. When one groups behavioral characteristics into a profile, the temptation is often clinically irresistible to elevate the profile to a syndrome and to infer that such profiles have a clear causal mechanism when they do not.

Even simple profiles, which are not labeled as syndromes but are simply behavioral descriptors, have been falling into legal disrepute for the same reasons. For example, in *United States v. Russell Banks* (1992) (25), a criminal conviction of CSA was overturned on the grounds that the MMPI did not provide a sufficiently sensitive profile of a "typical" child molester. That is, many people may have this personality test profile and not engage in sexually abusive behavior toward children. Given *Hadden* (19) and *United States v. Russell Banks* (25), behavioral profiles are legally suspect even though one could make a reasonable factual argument that doubt was significantly reduced by their use. The difference in the author's view is in the term *clinical utility*. For example, a given drug may significantly reduce average symptoms in a given sample of patients but if this effect is due to solid improvements in only a small proportion of these patients, then the drug has little practical use. Clearly then, we need to know something more relevant and specific about the patient before prescribing the drug.

Regarding CSA, we also clearly need to know something else about an alleged victim, accused or a claimant before useful behavioral profiles can be developed. For example, the measurement of risk of repeat offender violence (i.e., recidivism) has been recently achieved through the use of a reliable and valid behavioral profile called the Violence Risk Appraisal Guide (26), which includes items such as school problems, offending prior to age 16, being male, never married, a history of substance abuse, and a high

score on the Psychopathy Checklist-Revised. But this psychometric profile is numerically valid only when used with prior violent offenders. Due to base rate problems, it has no established predictive validity for nonoffenders. For the same reason, the Sexual Offense Recidivism Guide (26) is only useful for predicting sexual offenses for those with a record of sexual offenses. In terms of CSA profiles, we need to know which child victims have preexisting conditions or areas of vulnerability, such as low self-esteem or internalizing disorders, so that the content and intensity of current presenting problems attributed to CSA can be logically connected to the content of these prior conditions. Absent such information, how can we correctly attribute symptoms to CSA?

In Hadden, the syndrome at issue was CSAAS (21,22). This syndrome consisted essentially of sexualized behavior, behavioral reactions such as aggression or under-achievement, emotional reactions such as sleep disturbances and depression, and delayed reporting, and has no established predictive validity. The appellant successfully argued in this case that because psychologists' non-testamentary opinions are usually admissible in court, they are cloaked in an aura of *scientific infallibility*. The jury will then naturally assume that there is an acceptable scientific basis to these opinions even when there may be none—or it may be too weak. Lawyers may fail to cross-examine the expert witness (Rule 705), and therefore the jury and judge may be improperly misled. Unvalidated behavioral profiles are legal Trojan horses. According to Hadden, as novel and scientific evidence, syndrome testimony must pass the Frye test before being presented to the jury. Because CSAAS is the subject of considerable controversy (27), it did not pass this test. Indeed Summit (22) admitted that CSAAS was never designed to be diagnostic but is better suited to describing children whose diagnosis of sexual abuse has already been established. Nevertheless, he still vigorously defends its use in court.

The assessment of *psychic injury* is particularly challenging since CSA is often accompanied by the victim's feelings of shame (11). Complicating matters is its frequent use in civil proceedings by adult petitioners as a basis for punitive damage claims stemming from their proven or alleged CSA victimology as minor children. While sympathizing profoundly with the personal sufferings of victims, malingering is a very real concern when large sums of money are involved. Since a history of CSA is usually not accompanied by later adult mental health problems (28), the appeal to psychological injury syndrome evidence such as "the classic pattern of male sex abuse victims" in *Diocese of Winona v. Interstate Fire & Cas. Co.* (1994) (29) is scientifically suspect. In this proceeding, the proven victim was described as feeling worthless, shameful, unworthy, unable to link love with intimacy, depression prone, powerless, and distrusting.

These psychic injury cases raise profound forensic and legal questions. For instance, while convicted offenders can be held responsible for the natural consequences of their acts, can they also be held responsible for victims' pre-existing symptoms that are misattributed to the effects of their sexual crimes? It is necessary to disentangle such preexisting conditions from those which may result from CSA. Because the effects of CSA may have no definably unique features, the intensity and scope of the victim's preexisting conditions may be exacerbated by the abuse. Another useful question for forensic research is to predict which CSA victims are more vulnerable to long-term subjective distress. That is, do CSA victims who have long-term emotional symptoms also suffer from pre-existing emotional conditions that place them at a relatively higher risk of having these problems magnified by CSA? Also, are there some sexually abusive behaviors that are relatively more

likely to give rise to long-term emotional injuries when combined with preexisting emotional vulnerabilities? For example, forcible sexual abuse with penetration during early adolescence seems to place boys at a relatively higher risk for long-term problems (28). Are such long-term problems more predictable when the adolescent has preexisting family problems and low self-esteem? The comorbidity issue of CSA and lifestyle symptoms is discussed below.

The Empirical Problem with CSA Profiles

While certain behavioral, cognitive and emotional characteristics, such as sexualized behaviors (30,31), sexualized behaviors with aggression (32), and violence-related post-traumatic stresses (33) are significantly more likely in sexually abused than nonabused children (21), many nonabused children can also manifest these characteristics and about one-third of abused children are asymptomatic (11,32,34). Even children with verified sexually transmitted diseases (STDs) are as likely as not to be behaviorally asymptomatic (35). While the above studies show that such behavioral problems are often elevated in sexually abused children, there is also an unacceptable degree of disconnectedness between these behaviors and verified sexual abuse.

Moreover, comorbidity of sexual abuse with physical abuse and neglect is quite high (32), thereby confounding putative causative mechanisms. Indeed, using a meta-analysis of 59 studies of college students who self-reported a history of CSA, Rind, Tomovitch, and Bauserman (28) found that family environment accounted for three times the current mental health adjustment variance as the history of sexual abuse did. Moreover, the effects of family environment on current adjustment were moderate, accounting for 16 to 17% of the variance. Moreover, Carrey, Butter, Persinger, and Bialik (36) found that children with verified seriously sexual (involving penetration) or physical abuses were more likely to have significantly depressed WISC-R verbal and full-scale IQ scores. On the other hand, while cognitive impairment was associated with sexual abuse, cognitive impairment was not necessarily indicative of sexual abuse. Similarly, Persinger (37) found that groups of neurologically traumatized, conduct disordered, sexually abused, and reference group children could not be specifically identified on the basis of their WISC-R subtest scores, cautioning clinicians not to implicitly conclude a presumed etiology on the basis of behavioral differences in diagnostic groups.

Inappropriate sexualized behaviors and PTSD follow similarly obscure etiological patterns. Even though inappropriate sexual behaviors were significantly elevated in a sample of sexually abused children (30), these behaviors may have stemmed from a variety of sources, such as a neglectful lifestyle (28,32), or perhaps by means of self discovery, other than necessarily through exposure to illegal sexual experiences. More practically, most sexually abused children did not exhibit inappropriate sexual behaviors. Kendall-Tackett et al.'s (31) review showed that sexualized behavior and PTSD were consistently more likely in sexually abused than nonabused children; however, the incidence of sexualized behavior was about 35% for the sexually abused group but 17% for the physically abused group. These figures imply an estimated predictive power of about 18%. Even if elevated sexualized behavior is stable over time (30), suggesting that sexualized behavior is apparent in some child victims and not others, the sufficient predictive power is barely 50%. These are hardly figures which would impress a jury. PTSD was also not confined to the sexually abused group but was an indication of a "generalized stressor" (p. 174). Moreover, Finkelhor (33) questioned the usefulness of PTSD in the assess-

ment of CSA as such abuse is often not experienced as dangerous or threatening. Even when using human figure drawings, Bruening, Wagner and Johnson (38) found that clinical indicators of emotional disturbance (e.g., shaded crotches) did not discriminate between sexually abused and clinically referred girls; in fact, such indicators were scored more often when the clinicians doing the scoring were led to believe that the drawings were made by girls drawing a man and pretending to be sexually abused. Are we to conclude that projective tests are useful only in confirming our generic prejudices (2)?

Accordingly, there is still no known reliable and valid diagnostic profile of sexual abuse which would satisfy most courts (27). Even medical evidence is less than compelling. Surprisingly, in a review of this literature, Weissman (39) found that published genitourinary and perianal anatomical data are not usually normed against crucial base-rate information. Thus, except in the face of concrete anatomical data such as fresh abrasions, semen traces, pregnancy and STDs, even medical examinations may not yield clear and convincing evidence of sexual abuse. For example, Heger (40) found that in almost half the cases, even sodomy may leave no physical signs of trauma. Quantitatively controlled studies are needed in this vital area.

Alternatives to Syndromes: Enter the Scientist-practitioner as Problem-solver and Educator

Clinical training and tests typically do not address psycholegal problems directly and as such it is the assessor's use of conceptual validity that sets a good forensic assessment apart from a bad one (41). Conceptual validity involves addressing hypotheses about an individual using appropriate data toward the construction of a model about that individual which helps to explain their behavior. This process is useful and valid to the court to the extent that certain competing hypotheses are weighed and observations flow naturally from the model which is most consistent with the facts. A useful role for the psychologist investigating CSA is to form a working model of children's functioning before and after the alleged offense(s) and compare it to other alternatives (e.g., family conflicts, being bullied, having a serious accident, a recent move, etc.) which may as easily account for any significant behavioral changes. It is just such an approach that is used in preparing assessments of personal injury and claims of diminished capacity (41). In evaluating CSA, the forensic assessor must be sensitive to the fact that CSA is not a definable psychological disorder but ". . . a heterogeneous group of behaviors perpetrated on children by offenders" (42, p. 9) which, because they are nonspecific, should not be confused with the offender's behavior. Fischer and Whiting (42) underscored the fact that forensic child evaluations that confine themselves to sound measures of children's psychological statuses are vital facts which allow judges and juries to evaluate competing hypotheses about a defendant's guilt or innocence.

A forensic practitioner may also serve as an expert advisor or *amicus curiae* to the court without conducting a forensic assessment. Evans (8) also provides for such an educational role in custody and access assessments in Ontario. This educational role of the expert witness is amply illustrated in *Bighead v. United States of America* (1997) (43), wherein a charge of CSA was affirmed because an expert's testimony citing the CSAAS profile was "exquisitely useful" to the jury in assisting them with their interpretation of key facts, such as the child's delayed reporting. As the reader would expect, the defense provided an eloquent and detailed objection to this evidence using Frye, Daubert, and the Federal

Rules of Evidence, challenging the expert's credentials and the subjective nature of the expert's testimony. However, the expert did not evaluate the child but merely assisted the jury in an educational fashion based on her extensive therapeutic experience with children who were suspected or known victims of sexual abuse. [If the expert advisor were to step outside of this educational role and make statements about the alleged child victim, then it would be a serious breach of ethics not to have evaluated her (43).] In *Bighead*, there was no danger of applying reverse logic as there was no causal syndrome inference being made here. It is not an error of fact to claim that CSAAS behaviors are common in victims of sexual abuse. The *Bighead* transcript was curiously silent as to whether the jury was properly educated as to the scientific limitations of CSAAS.

Children's Interviews as Evidence

Since there is currently no valid syndrome for CSA and there is a strong likelihood of CSA symptoms having multiple explanations, short of a valid offender confession (45) and incontrovertible medical evidence, the crucial aspect of a proper forensic evaluation of suspected sexual abuse in children is usually the interview itself, which must be done in an "ethically permissible manner" (16, p. 139). Mandatory state and provincial reporting laws require the health care professional as well as others in positions of trust to report the suspected sexual abuse of children to child protection services or the police. However, since the initial disclosure interview with the child is usually the most valid one (46), the clinical practitioner should be prepared to provide minimally optimal standards of forensic interviewing at this crucial juncture. Given the harmful effects of sexual abuse, failure to deal effectively with the child's disclosure is ultimately unsupportive.

Child protection workers, police, psychologists, psychiatrists, social workers, and lawyers are often unevenly trained in CSA interviews, and often labor under scientifically unfounded but strongly held CSA beliefs (16,46). It is, therefore, essential that all professionals working with children and families should make themselves acquainted with competent CSA interviews based on the free recall guidelines of statement validity analysis (SVA) (46), which are reviewed and updated below. We know much more about what harms the validity of an interview than what supports it, so it is important that CSA interview procedures avoid being sources of contamination so that Frye and Daubert challenges can be more effectively met.

Statement Validity Analysis (46) is a serious attempt to construct a valid and systematic forensic interview schedule for children that emphasizes free recall as its key component. Geiselman and Fisher's Cognitive Interview (47) is structured similarly to SVA except that specific memory enhancing instructions (e.g., reconstruct the circumstances, be complete, recall in different orders, change perspectives) are employed prior to the open narrative section and may be inappropriate for children less than seven years of age (11). It is important to note that SVA also employs 19 content criteria for statement analysis of initial interviews covering general characteristics (e.g., logical structure), specific content (e.g., contextual embedding), peculiarities of content (e.g., unusual details), motivation-related contents (e.g., spontaneous corrections), and offense-specific elements (e.g., details characteristic to the offense) which are similar in many respects to those included in Stellar and Koehnken's (48) criteria-based statement analysis. However, short of the broader free recall guidelines, at this time there is little or no empirical support (49) for SVA criteria dealing with the content

analysis of young children's statements or of the use of counter-suggestion. But free recall is easier to achieve in theory than in practice. The following tentative suggestions are offered.

Tentative CSA Interview Guidelines and Working Assumptions

Consistency is Not a Valid Indication of Children's Statement Accuracy—Judges may exercise consistency in establishing the reliability and necessity of a child's statements in *voir dire* before admitting them into evidence. [The reader is invited to read Regina v. Dubois, (50), for an exemplary explanation of this process by the presiding magistrate.] Alarming, psychological studies now cast doubt on the use of this sage legal benchmark, at least with young children. For example, Bruck et al. (16) examined SVA-like linguistic markers, such as the amount of information provided by the child, number of spontaneous unprompted statements, and coherency of narrative in children's statements following the Sam Stone classroom simulation (51). They found that when young children were repeatedly and suggestively interviewed, the characteristic of broad consistency was the single most valuable one of a true statement, although there were as many inconsistencies and consistencies within the details of both true and false statements. Countersuggestion may lead to inconsistencies which are products of the interview and do not necessarily indicate a false report. However, it is salutary to note that it is easier to change than to plant or erase children's memory for being touched (52).

Even under mathematically tight experimental conditions addressing fuzzy trace theory, Brainerd, Stein, and Reyna (53) found that 7- and 10-year-olds' explicit or conscious memory for words and nonsense words was better than their unconscious or gist memory of related distractors. However, while unconscious memory was significantly better than chance, it was significantly more prone to error. Moreover, the likelihood of error for both conscious and gist memory was greater for the younger children. Memory was substantially better when "primed" or preceded by the instantiating target, but when the prompting was misleading, important details of the memory were more inaccurate. Brainerd and Mojardin (54) also examined the impact of memory inoculation and mere testing on everyday verbal stimuli (e.g., a pear is harder than a peach) that have forensic appeal. Using 6-, 8-, and 11-year-olds and adults, these data showed that false alarm gist memories for related distractors were affected more than were hits and were more persistent over trials. First-graders were especially more likely to accept old words with new meanings and to therefore have the greatest persistence of false alarms for distractors of greatest forensic interest. Thus, the use of consistency for remembered events is not a valid marker of their truth value, especially for young children. Appearance-reality distinctions (e.g., "looks like") are particularly difficult for children from 6 to 7 years of age or under (55).

Keep Interviews to a Minimum—Ideally, when allowed to run its natural unfettered course, the initial disclosure is the most valid basis of a CAS interview. As the number of interviews and people involved in doing them increase (3,16,46), contaminating influences also increase. Unfortunately, according to Van De Kamp (56), it is common for children to be interviewed on an average of seven times by police and, as surveyed by Leichtman and Ceci (51), from 4 to 11 times plus several nonforensic interviews by families and other interested parties, thus seriously reducing interview validity, hence the need to properly deal with the initial disclosure. The empirical CSA precautions in this paper stand in sharp contrast to published police guidelines. For example, Goldstein (57) stated the fol-

lowing: "The more often the abuse occurred, the greater the number of interviews that will be needed" (p. 180).

Children Seldom Recant Unless They are Provoked to Do So—Bruck et al.'s (16) analysis of the recantation literature (e.g., 5) shows that the large majority of children with validated claims do not retract them when interviewed. Overly aggressive interviewing may actually lead the child to recant (11), as well as to create believable but false accounts (16). The risk of sexually abused children to recant ". . . may well live on as a stubborn urban legend among frontline workers" (16, p. 138).

Interview Children Without Their Parent(s) Present—All too often, the validity of the initial interview may be compromised by well-meaning parents, teachers, or healthcare practitioners whose initial involuntary reaction is one of high emotion, either as an ill-guided but well-intentioned way of alleviating their own and the child's distress, or to prevent retraction. Of course, it is quite important for the child to feel believed (12), but the parent's presence may exert undue pressure on the child. The child may say what the parent wants to say, not what the interviewer should hear.

It should not be assumed that a child's reluctance to discuss necessarily signifies anything more sinister about the seriousness of the alleged abuse as seriousness is not correlated with the child's decision to disclose. Sauzier (12) found that most (54%) children subjected to intercourse with aggression never revealed, but about as many (50%) never told after exhibitionism or voyeurism. Also, it was more usual (33 to 40%) for children to tell immediately after relatively minor forms of abuse, such as fondling or attempted sexual contact as opposed to intercourse (23%) or penetration (17%). Notwithstanding child protection issues, the short- and long-term repercussions to the child for disclosing are not necessarily therapeutic (12). In incestuous circumstances, it is generally more fortuitous for allegations to stem from independent sources as this reduces children's misperceptions of their own feelings of culpability in disrupting their families. Moreover, the clinical picture for disclosing children who have been abused is not significantly different from those who have been abused and who do not disclose (12).

Avoid Inappropriate Comments—Overly supportive reactions may be felt to be in keeping with good caretaker standards but are not commensurate with forensic ones (58). Naively unrealistic comments such as "Now everything is going to be alright" are quite unrealistic in light of the actual impact of disclosure on the child's life and may actually lead some children to retract (11,21). Children and adolescents are often either very ambivalent about disclosure and have either very unrealistic or painfully accurate perceptions of the repercussions to themselves and their families (12). For example, the author witnessed a 3-year-old in a botched CSA interview say, "Is my daddy gonna go to jail?"

Interviewing children about potential abuse is akin to Heisenberg's uncertainty principle, which suggests that we cannot effectively observe something without changing it in the process. While such changes can be systematically measured in the world of physics, in the relatively more mundane world of forensic interviews, they are legally problematic. Accordingly, legitimate efforts at rapport building should not give way to encouraging and reinforcing children's answers which do not foster accurate reporting (10). Bruck et al. (16) and Ceci and Bruck (10), in particular, cautioned against the use of aggressive and persistent lines of questioning (e.g., "You will feel better after you tell") and re-

peated directive probing which may create an atmosphere of accusation.

Establish Neutral Rapport—Suggestive and emotionally laden interviewing especially at this anxiety-laden juncture runs the real risk of filling the child’s memory in with emotional reinforcements, post event information and stereotypes which blur fantasy-reality distinctions and foster reconstructed memories for acts or words which may be factually inaccurate. Establishing a comfortable “nice guy” relationship with the child is far less likely to provide misleading abuse information (59).

Avoid Inducements—One should also not selectively reinforce children’s responses, either openly (e.g., “Wow, that’s great”), or by nonverbal gestures (e.g., nodding, smiling), or by bribing. Promises of reunification with their parents or that their daddy won’t go to jail, are invalidating, as are attempts to gain their cooperation by appealing to authority such as helping the police in an important case. In stark contrast, to gain suspected victims’ cooperation, Goldstein (57) advocated the use of badges made out of paper or cardboard for children to wear, so they can be deputized. This advice may seem appropriate in police circles but has no scientific basis and is contraindicated in sound forensic interviewing.

Children May Misidentify the Perpetrator—Proper identification of the alleged perpetrator is central to the CSA interview and may be overlooked even when interview accuracy is otherwise quite high. Goodman, Rudy, Bottoms, and Aman (60) conducted an innovative and ecologically valid experiment of children’s memory for abuse-related material. A staged event in a trailer was enacted by a confederate man engaging in abuse-related but innocuous behaviors (e.g., tickling) with each child. Even after 14 days, when leading questions were used to test 4- and 7-year-olds’ memory of the trailer event, even the younger children were very accurate in their recall. When commission errors did take place, there were no errors which related to abuse-related actions but a few 4-year-olds made errors (e.g., being kissed) typical of the kind which may lead to abuse investigations. However, the error rate for *perpetrator misidentification* (i.e., the trailer man) was 38% for the 7-year-olds and 61% for the 4-year-olds. The legal repercussions here are rather obvious.

Discuss Truth and Falsehood with the Child—Yuille (46) stressed this aspect of the CSA interview. Using children from 7 to 12 years of age and adults, Warren, Hulse-Trotter, and Tubbs (61) creatively improved their resistance to misleading and persistent questioning after being read a purse-snatching story. Interviewees who were given explicit and specific warnings that a few of the details in the post-event information were inaccurate had significantly higher correct recall. These results were also revealing as the younger participants were significantly more likely to change their stories in response to social pressure. Repeated questioning of children by high status adults has a detrimental effect on their testimony (16) and valid efforts to inoculate children against this are needed.

Children have Good Recall of Traumatic Events but CSA is not Necessarily Traumatic—Notwithstanding all the pitfalls of interviewing young children, it is helpful to note that preschoolers are often quite accurate in recalling traumatic events, such as a visit to the hospital to treat an injury (62). Peterson found that when responsibly interviewed after six months’ time, using questions an-

chored by prefixes such as when, where and who, young children over the age of three were 95% accurate in relating details about the causes of their injuries and the medical treatments they received. However, interviewing young children about sexual abuse is more difficult because the assumption of trauma does not necessarily hold (33) unless injury or aggression have been experienced and sexual abuse may also entail conflicting family loyalties (11). It will be recalled that most children experiencing intercourse with aggression never reveal (12) and sexual abuse is often not accompanied by danger, threat, and violence (12). The key issue for most disclosing children may then be their appreciation of a betrayal of trust and what this disclosure means to them and their families. This topic is deserving of considerably more research attention than it has otherwise attracted. Thus, remembering and reporting are not synonymous.

Avoid Using Anatomically Detailed Dolls—Yuille (46) advises the use of pictures only to assist preschoolers in naming body parts. However, great caution should be exercised here as well. Forensic assessors should refrain from using behavior with anatomically detailed dolls as valid CSA data (10,11). Bruck et al. (16) found that when 3- and 4-year-olds were provided with an anatomical doll and props, such as a stethoscope, misleading questions about a medical exam produced significantly and forensically telling errors. For example, a significant number of children who had not had a medical exam showed genital touching on the doll and a number of girls inaccurately showed the pediatrician inserting his finger or props into their vaginas or buttocks. The legal implications here are fairly obvious. Innocent people can and have been sentenced to death on the basis of children’s testimony (e.g., Texas vs. Macias, 1987, 63).

Believability is Not a Valid Criterion of a Credible Statement (16,51)—Even experts, such as research psychologists, judges, and social workers were found to be very inaccurate when judging children’s credibility on the basis of their videotaped reports (16). These experts were no better at identifying children who were relating a true account than a false one. The message here is clear. If experts can be so easily misled, then the trier of fact can be too. The assessor’s believability comments may then improperly invade the province of the jury.

A Sobering Lesson

Consider the extensive and misleading interviewing of an alleged preschool child victim of sexual assault in *R. v. Simpson* (1996, 64) in which a daycare worker’s conviction of CSA was overturned. In this case, the psychologist violated almost all of the above CSA interview guidelines. This child was suggestively and misleadingly interviewed for 50 or more sessions spanning over a year. The psychologist used multiple props, tangible reinforcements, educational videos and frequent references to the name of the accused in the context of the alleged offense. The original conviction was successfully appealed. The appeal was successful because the jury was not warned about the lack of factual basis of the expert testimony and because this expert opinion proceeded on the basis of naming the appellant, and therefore invading the province of the jury. One wonders what proportion of CSA proceedings in any given year adhere to a reasonable approximation of scientific standards in evaluating behavioral testimony. Let us hope that the proverbial scientific baby is not ultimately discarded along with the behavioral bathwater.

References

1. Finkelhor D, Dzuiba-Leatherman J. Victimization of children. *Am Psychol* 1994;49(3):173–83.
2. Vidmar N. Generic prejudice and the presumption of guilt in sex abuse trials. *Law and Hum Behav* 1997;21:5–25.
3. Coolbear JL. Credibility of young children in sexual abuse cases: assessment strategies of legal and human service professionals. *Can Psychol* 1992;33:151–64.
4. Corwin DL, Berliner L, Goddman J, White S. Child sex abuse and custody disputes: no easy answers. *J Interpersonal Violence* 1987;2:91–105.
5. Jones DPH, McGraw JM. Reliable and fictitious accounts of sexual abuse in children. *J Interpersonal Violence* 1987;2:27–45.
6. Raskin DC, Yuille JC. Problems of evaluating interviews of children in sexual abuse cases. In: Ceci SJ, Toglia MP, Ross DF, editors. *New perspectives on the child witness*. New York: Springer-Verlag, 1987.
7. Sink F. Studies of true and false allegations: a critical review. In: Nicholson EB, editor. *Sexual abuse allegations in custody and visitation cases*. Washington, DC: American Bar Association, 1988;37–47.
8. Evans DR. The law, standards of practice, and ethics in the practice of psychology. Toronto: Emond Montgomery Publications Limited, 1997.
9. Ontario Psychological Association. *Ethical guidelines for psychological practice related to child custody and access*. Toronto: Ontario Psychological Association, 1998.
10. Ceci SJ, Bruck M. Suggestibility of the child witness: an historical review and synthesis. *Psychol Bulletin* 1995;113:403–39.
11. Kuehnle K. *Assessing allegations of child sexual abuse*. Sarasota, FL: Professional Resource Press, 1996.
12. Sauzier M. Disclosure of child sexual abuse—for better or worse. *Psychiatr Clin North Am* 1989;12:455–69.
13. Federal/Provincial Task Force on Uniform Rules of Evidence. Canada. Uniform Law Conference of Canada. Report of the federal/provincial task force on uniform rules of evidence. Prepared for the Uniform Law Conference of Canada. Toronto: Carswell, 1982.
14. Cornell Law School. Legal Information Institute. Federal rules of evidence for United States courts and magistrates. [Web Page] 1999; www.law.cornell.edu/rules/fre/. [Accessed 25 May 1999].
15. Rotgers F, Barrett D. Daubert v. Merrell Dow and expert testimony by clinical psychologists: implications and recommendations for practice. *Prof Psychol* 1996;27:467–74.
16. Bruck M, Ceci SJ, Hembrooke H. Reliability and credibility of young children's reports: from research to policy and practice. *Am Psychol* 1998;53:136–51.
17. Ziskin J, Faust D. *Coping with psychiatric and psychological testimony*. Marina del Rey, CA.: Law and Psychology Press, 1988.
18. Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579, 589, 125 L. Ed. 2d 469, 113 S. Ct. 2786 (1993).
19. Hadden v. Florida, 690 So. 2d 573 (Fl. Sup. Ct. 1997).
20. Frye v. United States, 293 F. 1013 (D.C. App. 1923).
21. Summit R. The CSA accommodation syndrome. *Child Abuse and Neglect* 1983;7:177–93.
22. Summit R. Abuse of the CSA accommodation syndrome. *J Child Sexual Abuse* 1992;1:153–63.
23. Douglas J, Olshaker M. *Journey into darkness*. New York: Scribner, 1997.
24. Edwards CN. Behavior and the law reconsidered: psychological syndromes and profiles. *J Forensic Sci* 1998;43:141–50.
25. United States v. Russell Banks 36 M. J. 150 (CMA, 1992).
26. Quinsey VL, Harris GT, Rice ME, et al. *Violent offenders: appraising and managing risk*. Washington, DC: American Psychological Association, 1998.
27. Ceci SJ, Ross DF, Toglia M. Suggestibility of children's memory: psycho-legal implications. *J Exper Psychol: General* 1987;117:38–49.
28. Rind B, Tromovitch P, Bauserman R. A meta-analytical examination of assumed properties of child sexual abuse using college samples. *Psychol Bulletin* 1998;124:22–53.
29. Diocese of Winona v. Interstate Fire and Gas Company, 858 F. Supp 1407; U.S. Dist. Lexis 8555, 1994.
30. Friederich WN, Gambich P, Damon L, Hewitt SK, Koverola C, Lang RA, et al. The child sexual behavior inventory: normative and clinical comparisons. *Psychol Assess* 1992;4:303–11.
31. Kendall-Tackett KA, Williams LM, Finkelhor D. Impact of sexual abuse on children: a review and synthesis of recent empirical studies. *Psychol Bulletin* 1993;113:164–80.
32. Watkins B, Bentovim A. The sexual abuse of male children and adolescents: a review of current research. *J Child Psychol and Psychiatry* 1992;33:197–248.
33. Finkelhor D. Early and long term effects of child sexual abuse: an update. *Professional Psychology: Research and Practice* 1990;21:325–30.
34. Bonney-McCoy S, Finkelhor D. Psychosocial sequelae of violent victimization in a national youth sample. *J Consult and Clinical Psychol* 1995;63:726–36.
35. Lawson L, Chaffin M. False negatives in sexual abuse disclosure interviews. *J Interpersonal Violence* 1992;7:532–42.
36. Carrey NJ, Butter HJ, Persinger MA, Bialik RJ. Physiological and cognitive correlates of child abuse. *J Am Acad Child and Adoles Psychiatry* 1995;34:1067–75.
37. Persinger MA. A comment concerning confounding variables within "referral controls" for sexual/physical abuse studies [unpublished manuscript]. 1994 Clinical Neuropsychological Laboratory, Department of Psychology, Laurentian University, Sudbury, Ontario, Canada.
38. Bruening CC, Wagner WG, Johnson JT. Impact of rater knowledge on sexually abused and nonabused girls' scores on the Draw-A-Person: screening procedure for emotional disturbance (DAP:SPED). *J Personality Assess* 1997;68:665–77.
39. Weissman H. Forensic psychological examination of the child witness in case of alleged sexual abuse. *Am J Orthopsychiatry* 1991;61:48–58.
40. Heger AH. Child sexual abuse: the medical evaluation. In: Nicholson EB, editor. *Sexual abuse allegations in custody and visitation cases*. Washington, DC: American Bar Association, 1988.
41. Maloney MP. *A clinician's guide to forensic assessment*. New York: The Free Press, 1985.
42. Fisher CB, Whiting KA. The (mis)use of posttraumatic stress disorder to validate CSA. Register Report 1996;22:8–10.
43. Bighead v. United States, 128 F.3d 1329 (9th Cir. 1993).
44. American Psychological Association. *Ethical principles of psychologists and code of conduct*. Washington, DC: American Psychological Association, 1997.
45. Kassin SM. The psychology of confession evidence. *Am Psychol* 1997;52:221–33.
46. Yuille JC. The systematic assessment of children's testimony. *Can Psychol* 1988;29:247–62.
47. Geiselman RE, Fisher RP. The cognitive interview technique for victims and witnesses of crime. In: Raskin DC, editor. *Psychological methods in criminal investigation and evidence*. New York: Springer, 1989; 191–215.
48. Stellar M, Koehnken G. Criteria-based statement analysis. In: Raskin DE, editor. *Psychological methods in criminal investigation and evidence*. New York: Springer, 1989;217–45.
49. Ruby CL, Bringham JC. The usefulness of the criterion-based content analysis technique in distinguishing between truth and fabricated allegations. *Psychology, Public Policy, and Law* 1997;3(4):705–37.
50. Regina v. Dubois. (3d) 544, (Quebec C. A.).
51. Leichtman MD, Ceci S. The effects of stereotypes and suggestion on preschoolers' reports. *Develop Psychol* 1995;31:568–78.
52. Pezdek K, Roe C. The suggestibility of children's memory for being touched: planting, erasing, and changing memories. *Law and Human Behav* 1997;21(1):95–106.
53. Brainerd CJ, Stein LM, Reyna VF. On the development of the conscious and unconscious memory. *Develop Psychol* 1998;34:342–57.
54. Brainerd CJ, Mojardin AH. Children's and adults' spontaneous false memories: long-term persistence and mere-testing effects. *Child Develop* 1998;69:1361–77.
55. Flavell JH, Green FL, Flavell ER. Development of knowledge about appearance-reality distinction. With commentaries by M. W. Watson and J. C. Campione. *Monographs of the Society for Research in Child Development* 1986;51(212).
56. Van De Kampe JK. Report on the Kern County child abuse investigation. Office of the Attorney General of California, 1986.
57. Goldstein S. *The sexual exploitation of children: a practical guide to assessment, investigation, and intervention*. Boca Raton, New York: CRC Press, 1998.
58. Greenberg SA, Shuman DW. Irreconcilable conflict between therapeutic and professional roles. *Prof Psychol: Research and Practice* 1997; 28(1):50–7.
59. Maloney MP, presenter. CSA evaluation: research, assessment issues and legal applications [workshop sponsored by the American Academy of Forensic Psychology, Las Vegas, NV]. 1966 April.

60. Goodman GS; Rudy L, Bottoms B, et al. Children's concerns and memory: issues of ecological validity in the study of children's eyewitness testimony. In: Fivush, R, Hudson J. editors. *Knowing and remembering in young children*. New York: Cambridge University Press, 1990; 249–84.
61. Warren A, Hulse-Trotter K, Tubbs EC. Inducing resistance to suggestibility in children. *Law and Human Behav* 1991;15:273–85.
62. Peterson C. The preschool witness: errors in accounts of traumatic injury. *Can J Behav Sci* 1996;28:36–42.
63. *Texas v. Macias*, 733 S.W. 2d 192 Texas Crim. App.
64. *R. v. Simpson* 3d 461 Ontario C. A.

Additional information and reprint requests:
Brian J. Bigelow, Ph.D., C. Psych.
Laurentian University
Sudbury, Ontario, Canada
P3E 2C6