

CRITICAL REVIEW

Psychiatry & Behavioral Science

Use of statement validity analysis in minors alleging sexual assault: A systematic review

Emilie Wouters MSc¹  | Lauriane Constanty MSc^{1,2} | Sébastien Urben PhD² |
Joëlle Rosselet Amoussou Master IS³ | Jacques Gasser PhD⁴

¹Department of Psychiatry, Unit of Child and Adolescent Forensic Psychiatry, Lausanne University Hospital and University of Lausanne, Prilly, Switzerland

²Division of Child and Adolescent Psychiatry, Department of Psychiatry, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

³Medical Library-Cery, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

⁴Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland

Correspondence

Emilie Wouters, Department of Psychiatry, Unit of Child and Adolescent Forensic Psychiatry, Lausanne University Hospital and University of Lausanne, Bâtiment des Cèdres, 1, route de Cery, 1008 Prilly, Switzerland.
Email: emilie.wouters@chuv.ch

Abstract

This systematic review aims to report on the use of Statement Validity Analysis (SVA) with minors involved in criminal justice proceedings. We conducted a literature search of six bibliographic databases up to March 2024. Additional searches were performed using citation tracing strategies. Nineteen studies published between 1991 and 2023 were retained. Most were published between 1991 and 2000, mainly in the USA. A scientific gap was observed for 10 years before studies resumed between 2011 and 2022. These 19 studies involved 2931 children; most were girls ($n=2080$; 71%). The mean age was 9.4 years ($SD=2.40$; $min=2$; $max=17.5$). Most studies did not mention the nature of the relationship between the child and the alleged perpetrator of sexual violence, three studies involved intra-family violence and six studies involved victims of intra- and extra-family violence. Nearly 75% of the interviewers were trained with SVA methods. Most were mental health professionals (52.6%) or police officers (15.8%). No study used the SVA as a whole, 10 studies used 19 criteria of the Criteria-Based Content Analysis (CBCA), and no study used the Validity Check List (VCL). Most studies performed SVA on interview transcripts ($n=8$), and two studies performed their analysis on both verbatims and video. The conclusion of our literature review highlights the methodological weaknesses of these studies and encourages more research about the use of SVA in the judicial field to reduce the risk of misleading the judiciary.

KEYWORDS

adolescent, child and adolescent forensic psychology/psychiatry, forensic credibility assessments, statement validity analysis, victims of sexual violence

Highlights

- Use of SVA in minors alleging sexual assault.
- We identified 19 studies involved only CBCA, none with VCL.
- Misuse of the SVA in the criminal court may mislead the judiciary.
- This systematic review highlighted a scarcity of studies about utilization of SVA.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Journal of Forensic Sciences* published by Wiley Periodicals LLC on behalf of American Academy of Forensic Sciences.

1 | INTRODUCTION

The prevalence of sexual violence against children contrasts with the United Nations Convention on the Rights of the Child (1989), with 194 countries having ratified it (November 2009) and declaring that they are taking all legislative, administrative and social measures to protect children from sexual violence. The rate of sexual violence among girls worldwide is estimated at between 16.4% and 19.7%, and among boys at between 6.6% and 8.8%, affecting the lives of millions of children around the world [1], with potential psychological (e.g. lack of confidence, depression, acute stress disorders), physical (e.g., sleep disorders, delayed language and development) and/or social (e.g., withdrawal, social isolation) negative consequences. When a child or teenager alleges sexual violence, the speech collected during the police hearing may be subject to a credibility analysis required by the judicial authorities. In particular, in sexual assaults, the analysis of the credibility of children's statements often carries considerable weight in the criminal judgment. For instance, in some court cases, this analysis is used as evidence to justify the incarceration of the alleged perpetrator. This systematic review provides an understanding of the scientific evidence of the use of SVA in the legal context in order to better measure the risks of inducing miscarriages of justice. Sexual allegations made by minors are often part of court cases with no tangible evidence or witnesses, where the minor's word is confronted with that of the alleged perpetrator. The criteria used to validate an adult's testimony related to the structuring of his or her account: its clarity, chronology, intelligibility and coherence of the acts stated [2]. These criteria cannot be transposed to the testimony of minors, however, whose most credible statements are frequently disjointed, non-chronological, hesitant, marked by emotions, contradictions and retractions and with a vocabulary dependent on the age and development of the alleged victim [3]. In light of this disparity, analysis of the child's testimony can play a key role in the judicial process and in establishing the legal truth.

These expert reports, drawn up by forensic experts, aim to assess the credibility of the alleged child victim's account of sexual assault using the Statement Validity Analysis (SVA) tool [3, 4]. SVA has its roots in *Statement Reality Analysis* (SRA), created in 1954 and based on the premise that statements based on memories of real experiences differ qualitatively from fictitious or suggested memories [4]. This instrument has been the subject of little scientific study, however, and appears to be based more on clinical skills and practitioners' experience. The SVA tool, which originated in Sweden and Germany, was introduced in 1989 [5–7]. It should be noted that in Switzerland today, the SVA is the only tool recognized and recommended by the jurisprudence of the Federal Court (ATF 129 I 49, TF 6B_539/2010 of 30.05.2011) to carry out this type of expertise.

The SVA of the child's discourse is carried out in three stages [2]: the quality of the police hearing, the quantitative analysis of the *Criteria-Based Content Analysis* (CBCA) and the weighting by the *Validity Check List* (VCL). The CBCA consists of assessing how the child's words were collected in the semi-structured interview

established during the police interview to verify that this account was collected in a non-suggestive way, favoring spontaneous discourse. These police interviews are filmed to avoid re-interviewing the children several times and contaminating their testimony. This also avoids potentially traumatizing them by repeating the alleged facts. The children's testimony is then recorded on video. The interview is then transcribed verbatim, in other words, the transcription of all the words spoken by the child during his testimony and the questions she/he was asked. In many places, such as the United States, Europe and Israel, police interviews of minors are conducted according to the *National Institute of Child Health and Trauma Development* (NICHD); [8] developed to encourage free expression of narratives in minors, given their suggestibility and their stage of development [9]. With this protocol, trained police officers ask open-ended and avoid suggestive questions, enabling a greater amount of spontaneous information to be obtained [10] than when they ask questions subjectively [8]. A better quality of police interviewing favors an increase in the detail obtained in these statements [11]. The second stage of the SVA is carried out by a professional, forensic psychologist or psychiatrist, trained in the use of the tool on the basis of the verbatim and the viewing of the video of the police interview. This recording is used to analyze the minor's speech using an evaluation scale known as CBCA. This consists of 19 quantitative spontaneity criteria divided into five categories: (a) general characteristics of the statement, (b) specific characteristics, (c) particularities of the content, (d) content relating to the motivations for the statement and (e) the specific element of the offense. The number of criteria used in the CBCA's quantitative analysis is a minimum of eight, including the first five. The third stage of the SVA is based on a checklist known as the VCL. It consists of evaluating the context of the child's story according to 18 qualitative criteria grouped into four categories (the child's behavior, characteristics of the interview, considerations regarding the motives for disclosure and other evidence). According to Yuille [12], the VCL is therefore used to evaluate all aspects of the testimony apart from the content. The conclusion of the evaluation of the credibility of the child's speech must be based on taking into account the quality of the interview with the police, the number of criteria retained in the quantitative analysis of the CBCA and the weighting by the VCL to be valid. The overall SVA rating is used to determine whether or not the child's testimony is credible. Note that Melkman, Hershkowitz [13] observe that, in cases of sexual violence against minors, forensic experts are more likely to conclude that children whose accounts involve extra-familial aggressors are credible than those involving aggressors who are the child's parents.

The aim of this systematic review is to identify the extent of scientific research using SVA with minors who have alleged judicialized sexual violence (i.e., sexual violence prosecuted in criminal proceedings). Experimental research conducted in a laboratory context will not be examined in this systematic review and has already been the subject of literature reviews [14, 15]. On the other hand, this research aims to assess whether the use of SVA in a judicial context meets the required standards [5, 7, 16], such as the use of the 19 CBCA criteria and weighting using the 18 VCL criteria. Secondly, exogenous

factors, such as clinical status, the relationship between the child and the alleged perpetrator and the profile of the children concerned by these legal measures [17], as well as the level of qualification and training of SVA examiners, is also investigated. These elements are likely to influence the discourse of children who, in addition to being highly suggestible, are likely to suffer from post-traumatic stress. Indeed, Koss and colleagues [18] point out that memories linked to sexual violence may be less accessible than memories linked to other experiences with traumatic potential due to the state of dissociation frequently observed in victims after sexual violence.

2 | METHODS

2.1 | Procedure

The systematic review is reported according to the PRISMA 2020 guidelines [19]. The study protocol is available on PROSPERO (#CRD42022293087). The inclusion criteria for the studies are (1) a population of minors under 18 years of age, (2) victims of judicialized sexual violence, (3) minors interviewed for a credibility analysis (according to the SVA or CBCA). Studies targeting an adult population, a credibility analysis not including SVA or with allegations of physical abuse alone and experimental studies were not included. Two independent reviewers (EW and LC) conducted all stages of the process (from selection to data extraction). Consensus was reached by exchanging arguments between the two reviewers (i.e., psychologists specializing in forensic psychology and trained in the use of the SVA) when results differed from one another.

2.2 | Information sources and search strategy

The comprehensive search was conducted on March 2024, in collaboration with a medical librarian (JRA), in six bibliographic databases, Medline ALL Ovid, Embase.com, APA PsycInfo Ovid, Web of Science Core Collection, Cochrane Database of Systematic Reviews and Cochrane Central Register of Controlled Trials. The searches were performed without language or date restrictions. Backward citation search (performed manually, EW and LC) and forward citation search (performed with Citationchaser ([20], JRA)), based on studies that were included, did not find any new records of interest. Appendix S1 provides details regarding the search techniques, keywords and index terms used.

2.3 | Selection process

A total of 353 articles were extracted from the databases. Retrieved records were imported into Endnote 20 (Clarivate Analytics) and duplicates removed manually (JRA). 227 articles were screened and full-text review processes were performed by two independent reviewers using Rayyan. Disagreements were resolved through

consensus. Then, based on the title and abstract of the article, 45 studies were identified, but 19 studies were selected on the basis of a full reading of the article and corresponding to the inclusion criteria presented in the present protocol. Figure 1 provides the flow diagram.

2.4 | Risk of bias assessment

To assess possible article biases (such as participant selection, sample size, statistical methods or ethical considerations) and thus the quality of the selected studies, we used the Appraisal tool for Cross Sectional Studies (AXIS) analysis grid [21]. The AXIS is a critical appraisal tool focusing on study design and the quality of study design. The AXIS tool has been developed so that it can be used across disciplines to facilitate the inclusion of cross-sectional studies in systematic reviews. Through a series of 20 items, this instrument assessed the quality of the studies, particularly objectives, methodology, results, source of funding and ethical aspects.

3 | RESULTS

3.1 | Description

The AXIS instrument did not highlight any major flaws in the quality of the articles included (see Table 1), apart from a single study with several flaws [22].

Nevertheless, for six studies, the statistical methods were not sufficiently described to be repeatable. Moreover, study limitations were discussed in only 13 articles, and only three studies indicated that participant consent had been obtained.

The 19 studies included (see Table 2 and Table 3) in the present research concerned a total of 2931 children, most of whom were girls ($n=2080$; 71%). The mean age was 9.4 years ($SD=2.40$; $min=2$; $max=17.5$). It should also be noted that 73.7% of the interviewees were trained or briefed in the use of SVA, and most were mental health professionals (52.6%) or police officers/investigators (15.8%). Ten studies were published between 1991 and 2000 (52.6%), only one between 2001 and 2010 (5.3%), and eight between 2011 and 2020 (42.1%).

Regarding the number of CBCA criteria, most studies used all 19 ($k=10$) or 14 ($k=6$) of the 19 criteria defined in the CBCA. A study conducted in South Korea used only four CBCA criteria. Finally, none of the selected studies reported using the VCL. Regarding the relationship between the presumed child victim and the perpetrator of sexual violence, most studies did not specify the nature of the relationship ($k=9$), while three studies were carried out on child victims of intra-family violence and six studies included, in a mixed manner, children presumed to be victims of intra- or extra-family violence. Three of the last six studies specified that they included unknown data. Regarding the narratives used in the discourse credibility analysis, most studies relied on the interview transcripts ($k=8$).

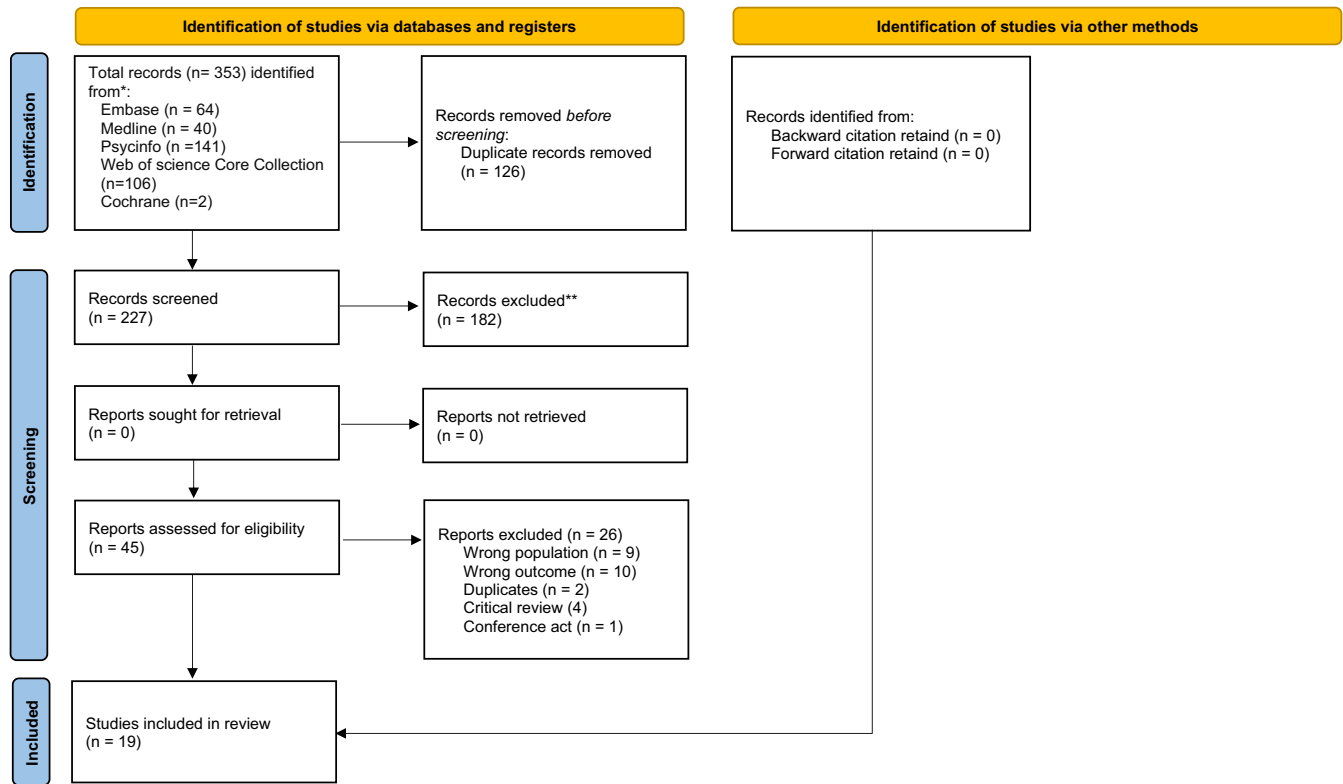


FIGURE 1 PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other sources. From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. <https://doi.org/10.1136/bmj.n71>. For more information, visit: <http://www.prisma-statement.org/>.

Only two studies carried out their analyses on verbatims and video-filmed interviews.

4 | DISCUSSION

We reviewed studies carried out in the field to report on the state of knowledge in favor of expert practices in a judicial context. The results showed that the state of knowledge is based on a small number of studies ($k=19$), mainly from the USA and Europe, carried out between 1991 and 2000.

We examined the socio-demographic characteristics of the samples studied, the material used to analyze the minors' speech, the number of CBCA criteria used, the use of VCL in the research, the training of the investigators, the socio-demographic characteristics of the children, information concerning their clinical status and the perpetrator-victim relationship in which the allegations were supported.

4.1 | Socio-demographic characteristics of the samples

Research into SVA with court-ordered minors began after the ratification of the International Convention on the Rights of the Child

(1989), suggesting that the needs of the judiciary seem to have been the driving force behind this field of study, and probably not the human sciences' knowledge of child development or language. In fact, SVA, which grew out of SRA, was formalized in 1989, at the same time as the Convention on the Rights of the Child (CRC) was adopted in New York in 1989 (RO 1998 2055). This temporal coincidence leads us to hypothesize that the evolution of children's rights in judicial matters have encouraged the publication of articles on the evaluation of children's speech in a judicial context using SVA. Subsequently, very few studies were carried out between 2001 and 2010. A slight rebound was observed between 2011 and 2017, where only two studies have been published on the subject, even though the tool is still used today in criminal courts. Since 2002, there have been no publications on the subject in the USA. This absence of scientific articles can be explained by the application of the Daubert (Daubert v. Merrell Dow Pharmaceuticals, Inc. 113 S. Ct. 2786, 1993) standards, which prevail in the courts of several American states. Indeed, the US Supreme Court has established strict conditions for the admissibility of expert opinions used as scientific evidence in court such as the validity of the expert's theory must be subject to research, the relevant evaluation and publications must be documented by peers, the error rate must be known and the scientific theory must be widely accepted by a scientific community and deemed reliable. We note that, according to Vrij [15], SVA analysis does not meet

TABLE 1 Bias assessments.

Authors	Year	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20
Akehurst et al.	2011	Yes	Yes	Yes	Yes	Don't know	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Don't know	Don't know
Anson et al.	1993	Yes	Yes	Yes	Yes	Don't know	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Don't know	Don't know
Boychuk	1991	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Don't know	Don't know
Buck et al.	2002	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Don't know	Don't know
Bybee, & Mowbray	1993	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Don't know	No	No	Yes	No	No	Yes	Don't know	Yes	Yes	Don't know	Don't know
Cabezas Garcia, et al.	2022	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Don't know	Yes
Craig, et al.	1999	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Don't know	Don't know
Davies et al.	2000	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Don't know	No	No	Yes	No	No	Yes	Don't know	Yes	Yes	Don't know	Don't know
Hershkowitz	1997	Yes	Yes	Yes	No	Don't know	Don't know	Don't know	Yes	Don't know	No	No	Yes	No	No	Yes	Don't know	Yes	No	Don't know	Don't know
Hershkowitz et al.	1999	Yes	Yes	Yes	Yes	Don't know	Yes	Don't know	Yes	Don't know	No	No	Yes	No	No	Yes	Don't know	Yes	No	Don't know	Don't know
Kim et al.	2011	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Don't know	Don't know
Lamb et al.	1997	Yes	Yes	Yes	Yes	Yes	Don't know	Don't know	Don't know	Don't know	No	No	Yes	No	Don't know	Yes	Don't know	Yes	No	Don't know	Don't know
Lamers-Winkelman & Buffing	1996	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Don't know	Don't know	No	No	No	No	No	Don't know	Don't know	No	No	Don't know	Don't know
Melkman et al.	2017	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	Don't know	Yes	Yes	Yes	Yes	Don't know	Don't know
Niveau et al.	2015	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	Don't know	Yes	Yes	Yes	Yes	Don't know	Don't know
Niveau	2021	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	Don't know	Yes	Yes	Yes	Yes	Don't know	Yes
Orbach & Lamb	1999	Yes	Yes	No	No	Don't know	Don't know	Don't know	Don't know	Don't know	No	Yes	Yes	Yes	Don't know	Yes	Don't know	Yes	No	Don't know	Don't know
Roma et al.	2011	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Don't know
Welle et al.	2016	Yes	Yes	Yes	Yes	Yes	Yes	Don't know	Yes	Yes	Yes	Yes	Yes	No	Don't know	Yes	Yes	Yes	Yes	No	Yes

Note: The list of items: Introduction: #1: Were the aims/objectives of the study clear?; Methods: #2: Was the study design appropriate for the stated aim(s); #3: Was the sample size justified?; #4: Was the target/reference population clearly defined? (Is it clear who the research was about?); #5: Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?; #6: Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?; #7: Were measures undertaken to address and categorize non-responders?; #8: Were the risk factor and outcome variables measured appropriate to the aims of the study?; #9: Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialed, piloted or published previously?; #10: Is it clear what was used to determine statistical significance and/or precision estimates? (e.g., p values, CIs); #11: Were the methods (including statistical methods) sufficiently described to enable them to be repeated?; Results: #12: Were the basic data adequately described?; #13: Does the response rate raise concerns about non-response bias?; #14: If appropriate, was information about non-responders described?; #15: Were the results internally consistent?; #16: Were the results for the analyses described in the methods, presented?; Discussion: #17: Were the authors' discussions and conclusions justified by the results?; #18: Were the limitations of the study discussed? Other: #19: Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?; #20: Was ethical approval or consent of participants attained?

TABLE 2 Included studies ($n=19$).

Author, publication date	Location	Sample size (N)	Girls and boys (N)	Age range	Formation	CBCA criteria	VCL
Akehurst et al. (2011)	UK	31	26/5	8–15	Yes	18	No
Anson et al. (1993)	USA	23	13/10	4.1–12.9	Incomplete	19	No
Boychuk (1991)	USA	75	60/15	NA	Yes	19	NA
Buck et al. (2002)	USA	104	80/24	2–14	Yes	19	No
Bybee, & Mowbray (1993)	USA	106	55/51	2.5–11	NA	19	No
Cabezas Garcia et al. (2022)	Spain	99	87 /12	4–17	Yes	NA	NA
Craig, et al. (1999)	USA	48	37/11	3–16	Incomplete	14	No
Davies et al. (2000)	UK	36	30/6	4–14	Yes	19	No
Hershkowitz (1997)	Israel	20	19/1	4–13	NA	14	NA
Hershkowitz et al. (1999)	Israel	24	NA	4–13	Yes	14	No
Kim et al. (2011)	Korea	189	163/26	8–13	Yes	4	No
Lamb et al. (1997)	USA	98	70/28	4–13	Yes	14	No
Lamers-Winkelmann & Buffing (1996)	Netherland	178	126/52	2–12	Yes	19	No
Melkman et al. (2017)	Israel	1563	1053/510	3–14	No	19	No
Niveau et al. (2015)	Switzerland	95	78/17	3–17	Yes	19	No
Niveau (2021)	Switzerland	96	50/46	6–16	Yes	19	No
Orbach & Lamb (1999)	USA	1	1/0	13	Yes	14	No
Roma et al. (2011)	Italy	109	86/23	4–14	Yes	14	No
Welle et al. (2016)	Switzerland	60	46/14	3.5–17.5	Yes	19	No

the Daubert guidelines set out above, given that the risk of error is 30% and that the SVA method is not unanimously accepted by the scientific community [15]. Amado [14], however, states that the risk of error was measured on the basis of laboratory studies but is actually 10% in field studies. He confirms that the tool's validity is recognized by the scientific community [14]. These arguments, therefore, support the degree of compliance with Daubert standards.

4.2 | Materials used to analyze the minors' discourse

We note heterogeneity in the material used to analyze children's speech. Four studies [23–26] guarantee that SVA analysis is based on verbatim (word-for-word transcription of children's narratives). Fifteen studies [13, 22, 27–39] analyze children's speech using transcribed interviews (i.e., report or summary of the child's testimony) or other materials without specifics or provide no information about how the child's speech was recorded and used. Almost a third of these studies do not guarantee that the SVA analysis was carried out on the basis of the verbatim: in this case, the speech delivered word for word by the minors. Furthermore, the studies do not specify whether there is any risk of contamination of the child's speech (i.e., the time between the allegation and the hearing, or the number of times the child repeated the alleged facts to a third party, or to how many people). It has been shown, however,

that conducting multiple interviews is problematic, as each interview offers an opportunity for suggestion, thereby diminishing the precision of the recall of facts [40]. Ceci and Huffman [41] note that suggestive information affects memory recall performance. Loftus and Davies [42] argue that, if an adult suggests information related to a disclosed event and the child has no knowledge to the contrary, she or he is likely to assimilate the adult's suggestion as fact. These authors underline the risk of contamination of children's discourse, given that they believe that adults have a better knowledge of the world than they do.

4.3 | Number of CBCA criteria

Initially, the CBCA was composed of 19 criteria, but six studies [22, 32, 34, 36–38] use 14 of them, systematically omitting the last five criteria (blank memory confession, doubts about one's own statement, disapproval of one's own participation, excusing the alleged perpetrator and the specific characteristic of the offense). The authors of these studies, however, offer no justification for their choices, other than the fact that they refer to an unpublished manuscript. The omission of items 15–18 may, however, alter the core meaning of the tool. It should also be noted that one study decided, without any precise rationale, to keep only four of the 19 criteria included in the CBCA, thus failing to meet the standards required for the use of the SVA and suggesting that misuse of the tool would render the results invalid.

TABLE 3 Descriptive.

		Total (k=19)	%
Origin	USA	7	36.8
	Israel	3	15.8
	Korea	1	5.3
	Switzerland	3	15.8
	Italy	1	5.3
	Netherland	1	5.3
	Spain	1	5.3
	England	2	10.5
Interviewer training	Yes	14	73.7
	No	1	5.3
	Partially	2	10.5
	NA	2	10.5
Material used	Verbatim	1	5.3
	Videos	2	10.5
	Verbatim and videos	2	10.5
	Verbatim and audios	1	5.3
	Transcribed interview	8	42.1
	Other	3	15.8
	NA	2	10.5
CBCA criteria	4	1	5.3
	14	6	31.6
	18	1	5.3
	19	10	52.6
	NA	1	5.3
Using the VCL	Yes	0	0
	No	16	84.2
	NA	3	15.8
Victim-aggressor relationship	Intra-family	3	15.8
	Extra-family	0	0
	Intra and extra family	7	36.8
	NA	9	47.4
Status	Victim	15	78.9
	Exposure	0	0
	Victim and + exposure	1	5.3
	NA	3	15.8

4.4 | Using the VCL

In our review of the literature, we note that none of the studies used weighting by VCL criteria. As a result, we note that the SVA, initially composed of three distinct stages, is not used in its entirety in all scientific research carried out in the field. Therefore, the results of

the VCL are not weighted by contextual factors that can modify the child's narrative (i.e., child's characteristics, how the interview is conducted, and considerations concerning the motives for disclosure or other evidence contained in the judicial file). The absence of the VCL does not rule out the risk of partial, erroneous or even invalid use of the SVA.

4.5 | Examiners training

Among the studies, we also note a wide disparity between the professions of the examiners analyzing minors' speech. Nevertheless, most articles indicate that mental health professionals initiate the assessment of children who allege sexual abuse. Among them, however, the training of those involved can vary widely (e.g., child psychiatrists, clinical psychologists, forensic psychologists or child/adolescent forensic psychiatrists). This heterogeneity of assessors makes it impossible to determine whether specific skills and knowledge regarding the child's clinical and developmental knowledge are required to use the SVA tool appropriately [43].

An analysis of the studies in this research shows that the training professionals receive on using the SVA tool is not very detailed and appears poorly structured. No inter-judge reliability is mentioned. In addition, very little information is available on the quality and duration of this training. These findings do not rule out the risk of intuitive or erroneous use of the SVA.

4.6 | Socio-demographic characteristics of minors

All the studies, apart from one case study, were based on a sample of predominantly girls rather than boys which is in line with the higher prevalence of girls reported sexually abused compared to boys [1, 44]. This gender disparity leads to the question of whether a boy's speech is treated or implicated differently in these analyses or in the court proceedings themselves. Further research on the subject would enable the further development of our hypotheses and knowledge. The results obtained in studies are not always consistent. Indeed, while several studies have found that more girls than boys report being victims of sexual assault [9, 45] other studies have found no gender-related difference in reporting [46, 47]. For boys who are sexually abused, it takes more than 10 years on average for them to report the abuse. For women, the average period between the experience of sexual violence and disclosure is shorter [48]. This hypothesis may help to explain the higher rates of girls reporting sexual violence than boys [1].

The age of the child who reports having been a victim of a sexual act significantly influences the characteristics of the report and, consequently, the CBCA score [49]. Up to the age of five, only 8% of children obtain the number of CBCA criteria required to conclude that the statement is credible [50]. In fact, cognitive and language development, as well as memory capacity, lead to both qualitative and quantitative improvement in the account given by the child

presumed to have been sexually assaulted, enabling a more detailed and contextualized account to be given. We note that 14 studies [13, 24, 26, 27, 29–32, 34–38, 51] mention using the SVA when the child is under 6 years of age, even though the tool is initially intended for minors aged six or over [2, 52, 53]. This raises the question of the scientific validity of using such a tool for younger children, whose cognitive, linguistic and emotional development is still in progress.

4.7 | Minors' status

Most studies ($k=16$) concerning minors who allege sexual assault emphasize that they are direct victims [13, 22–25, 27, 29–31, 33–39] rather than witnesses to such offenses. These studies do not, however, identify precise characteristics or information that would enable us to develop a profile or clinical status of the child. Apart from administrative information, such as age and gender, the studies provide no information on mental state, cognitive abilities or potential traumatic stress.

Indeed, a child's lexical field can influence his or her narrative: the younger the child, or the more intellectually delayed, the less able she/he is to provide details. The discourse of these children contains only the central action, resulting in a lower number of CBCA criteria. Craig, Scheibe [32] show that CBCA criteria are comparatively more numerous in the 10–16 age group than in the 3–9 age group. Davies, Westcott [35] observe that the number of CBCA criteria increases with age. Melkman, Hershkowitz [13] conclude their analysis by arguing that age, together with the absence of mental delay, was the strongest predictor of credibility.

Post-traumatic stress disorder (PTSD), for its part, can impair memory and recall in the context of sexual violence due to the dissociative state that ensues [18]. However, the presence or absence of a specific diagnosis of PTSD must be qualified, as the symptoms do not always appear after the event. In the case of minors interviewed shortly after an alleged sexual assault, enough time may not have elapsed for the diagnostic criteria for PTSD to have been met, and acute stress disorders or adjustment disorders are more likely to be observed. The absence of information on a possible traumatic-stress state means that we cannot exclude the influence of dissociation on the child's narrative, which also directly affects the number of CBCA criteria found in his or her speech. Moreover, we have no information about the frequency of alleged sexual assaults that might suggest symptoms specific to complex post-traumatic-stress states in situations of memory-altering sexual abuse (World Health Organization. ICD-1-Mortality and Morbidity Statistics, 2019) [54].

4.8 | Author–victim relationship

As for the context of the alleged sexual assault, this is generally not very detailed. While we can establish for some studies that the relationship between the child and the aggressor is intrafamilial ($k=3$, 15.8%) [22, 36, 37] and both intra- and extrafamilial ($k=6$, 31.6%)

[13, 23, 27, 28, 32, 38, 39]. Nevertheless, some data are missing from almost half of the selected studies ($k=9$, 47.4%) [24–26, 29–31, 33–35]. As a result, few studies reveal the nature of the relationship between the child and the alleged perpetrator.

Barudy [55] notes that the disclosure process and traumatic symptomatology depend on the nature of the relationship in which the sexual violence takes place. In the case of intra-familial violence or incest, the severity of the violence is amplified by a relational dynamic based on trust, secrecy and the risk of being excluded from the family unit in the event of disclosure. Extra-familial sexual abuse rarely involves such relational confusion. These aspects merit further investigation in the context of SVA use.

4.9 | Future directions

In light of these findings, it would be timely to conduct further research into the use of the SVA tool in the judicial context with minors who allege acts of sexual violence: on the one hand to improve knowledge on the subject and on the other hand to improve expert practices with a view to better serving justice. With this in mind, a new study is currently underway on SVA in its entirety in the Swiss canton of Vaud.

It would also be instructive to be able to compare studies carried out on an experimental basis with those involving material derived from actual events, to better identify the knowledge that can be transposed from one field to another and the disparities between them. In addition, although SVA is the only analysis tool recognized by Swiss jurisprudence, it would be scientifically useful to explore whether other tools or methods have been developed for similar purposes. Legal interpretation and examination of case law on admissibility would also be an important contribution, particularly for the United States as well as to compare it internationally.

4.10 | Limitations

Throughout this systematic review of the literature, we have focused on studies with children actually considered to be potential victims who have supported forensic sexual allegations, excluding laboratory studies. When it comes to discourse analysis using SVA, most studies concern research carried out in the laboratory on fictitious, non-judicialized discourse. We hypothesize that this limitation directly raises an ethical question: how can we carry out scientific studies with child victims and protect them without risking amplifying the traumatic injury they suffer?

5 | CONCLUSION

The courts, for credibility assessment, require an evaluation of a child's speech alleging sexual violence. The scientific validity of using SVA with potentially victimized children remains very fragile, however, as no study using the SVA tool in its entirety has been identified.

Our conclusions are in line with those of Bradford [56], who argued that SVA is a promising procedure but that there is a lack of research on norms that take into account the child's age, gender or life experience, affecting these assessments through the knowledge and skills of the assessors. Unfortunately, these conclusions remain valid 30 years later.

To avoid inducing judicial errors, however, it would be necessary for this tool to be the subject of more scientific research to be more standardized with a population of children directly involved in these criminal proceedings. We would like to emphasize the great contrast between the preponderance of these expert assessments in trials and the scientific fragility of the tool in the judicial context.

In the United States, there are differing opinions regarding the recognition of SVA as a criterion of proof according to Daubert standards [14], suggesting that both magistrates and healthcare professionals should be cautious about using this method. In France, since the Outreau case, the first trial of which took place in Saint-Omer in 2004, credibility expertise has been banned from mandates (Report of the National Assembly n° 3125 made on behalf of the commission of inquiry charged with investigating the causes of the dysfunctions of justice in the Outreau affair and formulating proposals to avoid their recurrence). This raises the question of whether countries that do not require credibility assessments for sexual allegations made by children, in view of the limited scientific references available, are not leaving the assessment of the credibility and coherence of the child's statements to the subjectivity of the judge. Experimental research [15] has shown that the SVA tool is more reliable than the clinical assessment of health professionals.

In Switzerland, although it is the only tool recognized in this field by case law, SVA is currently used in criminal courts with children and adolescents presumed to be victims of sexual violence. This literature review, however, underlines the fact that scientific references are part of a non-linear evolution that needs to be updated and that no field study has yet been carried out using the SVA method in its entirety. This raises ethical questions about child protection when such studies are considered. We feel, however, that it is essential to develop new research on this tool so as to be able to bring scientific elements to questions that often remain eminently subjective. Indeed, the question of the credibility of allegations of child sexual abuse frequently poses difficulties for judges. Being able to rely on scientifically validated tools would help improve judicial procedures. Incorrect use of the SVA would distort the analysis of the child's testimony. In the judicial context, the major ethical risk is to mislead the judiciary, thus, preventing the search for the truth and the effectiveness of criminal justice that every citizen has the right to expect.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

ORCID

Emilie Wouters  <https://orcid.org/0000-0003-3177-4976>

REFERENCES

1. Stoltenborgh M, van Ijzendoorn MH, Euser EM, Bakermans-Kranenburg MJ. A global perspective on child sexual abuse: meta-analysis of prevalence around the world. *Child Maltreat*. 2011;16(2):79–101. <https://doi.org/10.1177/1077559511403920>
2. Wouters E, Gasser J. Actualité de l'expertise de crédibilité chez les mineurs [state of the art about credibility assessment of minors]. *Rev Med Suisse*. 2018;14(619):1651–4. <https://doi.org/10.53738/REVMED.2018.14.619.1651>
3. Bénézech M. Vérité et mensonge: L'évaluation de la crédibilité en psychiatrie légale et en pratique judiciaire [Truth and lies: Assessing credibility in forensic psychiatry and judicial practice]. *Ann Med Psychol (Paris)*. 2007;165(5):351–64. <https://doi.org/10.1016/j.amp.2007.03.011>
4. Undeutsch U. The development of statement reality analysis. In: Yuille JC, editor. *Credibility assessment*. Dordrecht, Netherlands: Springer Netherlands; 1989. p. 101–19. https://doi.org/10.1007/978-94-015-7856-1_6
5. Steller M, Koehnken G. Criteria-based statement analysis. In: Raskin DC, editor. *Psychological methods in criminal investigation and evidence*. New York, NY: Springer Publishing Company; 1989. p. 217–45.
6. Raskin DC, Esplin PW. Statement validity assessment: interview procedures and content analysis of children's statements of sexual abuse. *Behav Assess*. 1991;13(3):265–91.
7. Yuille JC. The systematic assessment of children's testimony. *Canadian Psychol*. 1988;29(3):247–62. <https://doi.org/10.1037/h0079769>
8. Cyr M. Recueillir la parole de l'enfant témoin ou victime. De la théorie à la pratique [gathering the word of the child witness or victim. From theory to practice]. Paris, France: Dunod; 2023.
9. Hershkowitz I, Lamb ME, Katz C. Allegation rates in forensic child abuse investigations: comparing the revised and standard NICHD protocols. *Psychol Public Policy Law*. 2014;20(3):336–44. <https://doi.org/10.1037/a0037391>
10. Benia LR, Hauck-Filho N, Dillenburg M, Stein LM. The NICHD investigative interview protocol: a meta-analytic review. *J Child Sex Abus*. 2015;24(3):259–79. <https://doi.org/10.1080/10538712.2015.1006749>
11. Milne R, Bull R. *Investigative interviewing: psychology and practice*. West Sussex, England: Wiley; 1999.
12. Yuille JC. L'entrevue de l'enfant dans un contexte d'investigation et l'évaluation systématique de sa déclaration [Interviewing children in an investigative context and systematically assessing their statements]. Paris, France: Editions du Meridien; 1992. p. 67–113.
13. Melkman EP, Hershkowitz I, Zur R. Credibility assessment in child sexual abuse investigations: a descriptive analysis. *Child Abuse Negl*. 2017;67:76–85. <https://doi.org/10.1016/j.chiabu.2017.01.027>
14. Amado BG, Arce R, Farina F. Undeutsch hypothesis and criteria based content analysis: a meta-analytic review. *Eur J Psychol Appl Legal Cont*. 2015;7(1):3–12. <https://doi.org/10.1016/j.ejpal.2014.11.002>
15. Vrij A. Criteria-based content analysis: a qualitative review of the first 37 studies. *Psychol Public Policy Law*. 2005;11(1):3–41. <https://doi.org/10.1037/1076-8971.11.1.3>
16. Raskin DC, Yuille JC. Problems in evaluating interviews of children in sexual abuse cases. In: Ceci SJ, Ross DF, Toglia MP, editors. *Perspectives on children's testimony*. New York, NY: Springer New York; 1989. p. 184–207. https://doi.org/10.1007/978-1-4613-8832-6_9
17. Gewehr E, Hensel B, Volbert R. Predicting disclosure latency in substantiated cases of child sexual abuse. *Child Abuse Negl*. 2021;122:105346. <https://doi.org/10.1016/j.chiabu.2021.105346>

18. Figueredo A, Bell I, Tharan M, Tromp S. Traumatic memory characteristics: a cross-validated mediational model of response to rape among employed women. *J Abnorm Psychol.* 1996;105:421–32. <https://doi.org/10.1037/0021-843X.105.3.421>
19. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *J Clin Epidemiol.* 2021;134:178–89. <https://doi.org/10.1016/j.jclinepi.2021.03.001>
20. Haddaway N, Grainger M, Gray C. Citationchaser: an R package and shiny app for forward and backward citations chasing in academic searching. *Zenodo.* 2021;16:10-5281.
21. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open.* 2016;6(12):e011458. <https://doi.org/10.1136/bmjopen-2016-011458>
22. Orbach Y, Lamb ME. Assessing the accuracy of a child's account of sexual abuse: a case study. *Child Abuse Negl.* 1999;23(1):91–8. [https://doi.org/10.1016/S0145-2134\(98\)00114-8](https://doi.org/10.1016/S0145-2134(98)00114-8)
23. Akehurst L, Manton S, Quandt S. Careful calculation or a leap of faith? A field study of the translation of CBCA ratings to final credibility judgements. *Appl Cogn Psychol.* 2011;25(2):236–43. <https://doi.org/10.1002/acp.1669>
24. Niveau G, Lacasa MJ, Berclaz M, Germond M. Inter-rater reliability of criteria-based content analysis of children's statements of abuse. *J Forensic Sci.* 2015;60(5):1247–52. <https://doi.org/10.1111/1556-4029.12816>
25. Niveau G. Sensory information in children's statements of sexual abuse. *Forensic Sci Res.* 2021;6(2):97–102. <https://doi.org/10.1080/20961790.2020.1814000>
26. Welle I, Berclaz M, Lacasa MJ, Niveau G. A call to improve the validity of criterion-based content analysis (CBCA): results from a field-based study including 60 children's statements of sexual abuse. *J Forensic Leg Med.* 2016;43:111–9. <https://doi.org/10.1016/j.jflm.2016.08.001>
27. Anson DA, Golding SL, Gully KJ. Child sexual abuse allegations: reliability of criteria-based content analysis. *Law Hum Behav.* 1993;17(3):331–41. <https://doi.org/10.1007/BF01044512>
28. Boychuk TD. Criteria-Based Content Analysis of Children's Statements About Sexual Abuse: A Field-Based Validation Study. *Dissertation Abstracts International Section A: Humanities and Social Sciences.* Arizona State University. 1991;52(6-A):2272.
29. Buck JA, Warren AR, Betman SI, Brigham JC. Age differences in criteria-based content analysis scores in typical child sexual abuse interviews. *J Appl Dev Psychol.* 2002;23(3):267–83. <https://doi.org/10.1016/S0193-3973%2802%2900107-7>
30. Bybee D, Mowbray CT. An analysis of allegations of sexual abuse in a multi-victim day-care center case. *Child Abuse Negl.* 1993;17(6):767–83. [https://doi.org/10.1016/S0145-2134\(08\)80007-5](https://doi.org/10.1016/S0145-2134(08)80007-5)
31. Cabezas-García M, Fernández MEP, Rodríguez JMA. Psychosocial factors associated with the assessment of testimony credibility in child sexual abuse. *Anales De Psicología.* 2022;38(2):307–15. <https://doi.org/10.6018/analesps.418161>
32. Craig RA, Scheibe R, Raskin DC, Kircher JC, Dodd DH. Interviewer questions and content analysis of children's statements of sexual abuse. *Appl Dev Sci.* 1999;3(2):77–85. https://doi.org/10.1207/s1532480xads0302_2
33. Kim TK, Choi S, Shin YJ. Psychosocial factors influencing competency of children's statements on sexual trauma. *Child Abuse Negl.* 2011;35(3):173–9. <https://doi.org/10.1016/j.chiabu.2010.09.005>
34. Roma P, Martini PS, Sabatello U, Tatarelli R, Ferracuti S. Validity of criteria-based content analysis (CBCA) at trial in free-narrative interviews. *Child Abuse Negl.* 2011;35(8):613–20. <https://doi.org/10.1016/j.chiabu.2011.04.004>
35. Davies GM, Westcott HL, Horan N. The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychol Crime Law.* 2000;6(2):81–97. <https://doi.org/10.1080/10683160008410834>
36. Hershkowitz I. The dynamics of interviews involving plausible and implausible allegations of child sexual abuse. *Appl Dev Sci.* 1999;3(2):86–91. https://doi.org/10.1207/s1532480xads0302_3
37. Hershkowitz I, Lamb ME, Sternberg KJ, Esplin PW. The relationships among interviewer utterance type, CBCA scores and the richness of children's responses. *Legal Criminol Psychol.* 1997;2(2):169–76. <https://doi.org/10.1111/j.2044-8333.1997.tb00341.x>
38. Lamb ME, Sternberg KJ, Esplin PW, Hershkowitz I, Orbach Y, Hovav M. Criterion-based content analysis: a field validation study. *Child Abuse Negl.* 1997;21(3):255–64. [https://doi.org/10.1016/S0145-2134\(96\)00170-6](https://doi.org/10.1016/S0145-2134(96)00170-6)
39. Lamers-Winkelmann F, Buffing F. Children's testimony in The Netherlands: a study of statement validity analysis. *Crim Justice Behav.* 1996;23(2):304–21. <https://doi.org/10.1177/0093854896023002004>
40. Moston S. The suggestibility of children in interview studies. *First Lang.* 1987;7(19):67–78. <https://doi.org/10.1177/014272378700701904>
41. Ceci SJ, Huffman MLC. How suggestible are preschool children? Cognitive and social factors. *J Am Acad Child Adolesc Psychiatry.* 1997;36(7):948–58. <https://doi.org/10.1097/00004583-199707000-00017>
42. Loftus E, Davies G. Distortions in the memory of children. *Aust J Soc Issues.* 2010;40:51–67. <https://doi.org/10.1111/j.1540-4560.1984.tb01093.x>
43. Cooper BS, Griesel D, Ternes M, editors. *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment.* New York, NY, US: Springer Science + Business Media; 2013. p. xxii, 349. <https://doi.org/10.1007/978-1-4614-5547-9.editors>
44. Pereda N, Guilera G, Fornis M, Gómez-Benito J. The prevalence of child sexual abuse in community and student samples: a meta-analysis. *Clin Psychol Rev.* 2009;29(4):328–38. <https://doi.org/10.1016/j.cpr.2009.02.007>
45. Lippert T, Cross TP, Jones L, Walsh W. Telling interviewers about sexual abuse predictors of child disclosure at forensic interviews. *Child Maltreat.* 2009;14(1):100–13. <https://doi.org/10.1177/1077559508318398>
46. Goodman-Brown TB, Edelstein RS, Goodman GS, Jones DPH, Gordon DS. Why children tell: a model of children's disclosure of sexual abuse. *Child Abuse Negl.* 2003;27(5):525–40. [https://doi.org/10.1016/S0145-2134\(03\)00037-1](https://doi.org/10.1016/S0145-2134(03)00037-1)
47. Leach C, Powell MB, Sharman SJ, Anglim J. The relationship between children's age and disclosures of sexual abuse during forensic interviews. *Child Maltreat.* 2017;22(1):79–88. <https://doi.org/10.1177/1077559516675723>
48. O'Leary PJ, Barber J. Gender differences in silencing following childhood sexual abuse. *J Child Sex Abus.* 2008;17(2):133–43. <https://doi.org/10.1080/10538710801916416>
49. Vrij A, Akehurst L, Soukara S, Bull R. Will the truth come out? The effect of deception, age, status, coaching, and social skills on CBCA scores. *Law Hum Behav.* 2002;26(3):261–83. <https://doi.org/10.1023/A:1015313120905>
50. Van Gijsegem H, Gauthier M-C. Links between sexual abuse in childhood and behavioural disorders in adolescent girls: a multivariate approach. *Can J Behav Sci.* 1994;26(3):339–52. <https://doi.org/10.1037/0008-400X.26.3.339>
51. Lamers-Winkelmann F, Buffing F. Children's testimony in The Netherlands. *Crim Justice Behav.* 2016;23(2):304–21. <https://doi.org/10.1177/0093854896023002004>
52. Vrij A, Akehurst L, Soukara S, Bull R. Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. *Hum*

- Commun Res. 2004;30(1):8–41. <https://doi.org/10.1111/j.1468-2958.2004.tb00723.x>
53. Niveau G, Berclaz M, Lacasa M-J, With S. Mise en oeuvre du protocole d'évaluation de crédibilité SVA dans le contexte médico-légal francophone [Implementation of the SVA credibility assessment protocol in the French forensic context]. *Schweiz Arch Neurol Psychiatr.* 2013;164(3):99–106.
54. Dubois VH, Tordjman S, Delcey G, Vabres N, Dayan J, Balençon M. Psychotraumatisme de l'enfant et de l'adolescent: Définitions, aspects cliniques et principes de soins [Child and adolescent psychotrauma: Definitions, clinical aspects and principles of care]. *Perfectionn en Pédiat.* 2019;2(2):152–62. <https://doi.org/10.1016/j.perped.2019.04.018>
55. Barudy J. Les abus sexuels extra et intrafamiliaux [Extra and intra-familial sexual abuse]. In: Érès, editor. *La douleur invisible de l'enfant [the invisible pain of the child]*. Toulouse, France: Érès; 2007. p. 161–82.
56. Bradford R. Developing an objective approach to assessing allegations of sexual abuse. *Child Abuse Rev.* 1994;3(2):93–101. <https://doi.org/10.1002/car.2380030206>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wouters E, Constanty L, Urben S, Amoussou JR, Gasser J. Use of statement validity analysis in minors alleging sexual assault: A systematic review. *J Forensic Sci.* 2024;00:1–11. <https://doi.org/10.1111/1556-4029.15604>