

An analysis of eyewitness report in children: Using the CBCA with 7-year-old children

Giuliana Mazzoni
Seton Hall University

Katia Ambrosio
Università della Calabria

On child sexual abuse

Sexual abuse refers to the act of performing sexual activities with minors or other people who are not able to make choices or who are not aware of the real meaning and implication of the actions (Montecchi, 1994). We cannot think of child sexual abuse as simply being a distorted relationship between a “bad” adult and a helpless child. Child sexual abuse is the result of more complex interactions between social and family factors. Focusing attention on just one of the factors would not facilitate understanding the complex interaction that leads to the violent act. The act is located within a family or social group to which the victim belongs (Tortolani, 1991). Cases are frequently made public because of sudden or dramatic changes in the family situation, e.g. adolescent rebellions, antisocial behaviors, pregnancy, a psychiatric illness. Sometimes even events as simple as a fight can trigger the public accusation of sexual abuse.

Many child sexual abuse cases are not solved because diagnosing sexual abuse can be extremely difficult. There might be lack of physical marks, the abusing person might display moral integrity, and the abuse might be performed in secrecy, without any witnesses other than the victim, the child might lack the verbal skills to narrate the episode(s), and people might doubt child verbal reports about sexual abuse (Montecchi, 1994). All of these factors often make it difficult to identify the existence of a traumatic experience in the child.

Physical, cognitive, and emotional signs can be used as criteria in ascertaining the occurrence of sexual abuse. The physical signs include: defloration, the tearing of the frenulum, symptoms of venereal diseases as well as others that are much less clear, such as bruises, scratches etc in the vaginal or anal areas; wounds in the mouth or throat; torn or stained underwear. The physical diagnosis of sexual presents special difficulty, because very often with no penetration there is no medical evidence.

Cognitive signs comprise a level of knowledge of sexual terms and activities advanced for the age (mimicking actions, using words, drawing sexual activities). Emotional signs include anxiety, sudden tears, insomnia, phobic symptoms (avoiding going outside), obsessive symptoms (cleaning rituals), hostility and aggression. These behaviors, however, are often present also in many cases that have nothing to do with sexual abuse, and therefore they cannot be used as clear diagnostic symptoms. In other cases, one can observe a sudden decline in school performance, problems in reasoning, depression, inappropriate sexual behaviors (excessive masturbation, showing the genitals, introduction of objects in the genitals, forcing peers into sexual activities, etc. (Cosentino, MeyerBahlburg, Alpert, Weinberg, Gaines, 1995; Mayall, Gold, 1995).

Abused adolescents frequently show excessive sexuality, gender confusion (particularly in males that were abused by adult males), and sexual acting out (Urquiza & Capra, 1990). Also observed are antisocial behavior, drug and alcohol abuse, running away from home, decline in self esteem (e.g. “I am a prostitute”), sudden mood changes, guilt feelings, impotence, and, in some extreme cases, suicidal tendencies.

None of the above signs (excluding some signs of physical trauma) can be used as confirmation of sexual abuse, because all of them can be produced by other causes.

Drawing strong conclusions from the presence of these signs would lead into an error in judgment called “illusory correlation” (Gulotta et al, 1996). In girls, for example, a diagnosis of neovascularization is caused also by simple vaginal infections. The lack of integrity of the hymen can also be caused by congenital errors. This ambiguity also characterizes the behavioral and cognitive signs. Nightmares, excessive masturbation, depression, extensive sexual knowledge (inappropriate for the age), are not clear unequivocal signs of sexual abuse. Extensive sexual knowledge can also be acquired through videos, movies, or watching adults talking or behaving (Mazzoni, 2001).

Problems with the interview: reliability and validity of eyewitness testimony

The accusation of sexual abuse starts a legal procedure aimed at ascertaining the existence of the abuse and the role of the alleged perpetrator. It is easy to see from the previous paragraph how difficult it is to reach a firm decision about the occurrence of sexual abuse. The signs are not always present, and, if present, they can be ambiguous; the adult usually demands that the victim keep the abuse secret; the perpetrator refuses to admit having committed the act. In this situation, the testimony of the victim is crucial. Often it provides the only basis for the accusation. However the child witness faces a new and shocking situation when giving his or her testimony. It is important to neutralize the risk of the interview being traumatic and the child not talking. The interview should be conducted by specialized personnel. Simply being a developmental psychologist is of little or no help here, and does not guarantee the quality of the testimonial report. The interviewer should have a specific training in investigative interviewing, the knowledge and ability to interact with policemen, lawyers, judges, foster parents, and the ability to follow the child up to the final phase of the procedure (Sergio, 1997). Whereas well done interviews can support the child and produce reliable reports, poorly done interviews can have devastating consequences on the child’s life and compromise his or her psychological development.

When poorly done, interviews may contain leading and misleading questions that contain suggestive information. The interviewer needs to avoid being intimidating, and implicitly or explicitly demanding that the child reveal the “terrible event” that happened to him or her. The initial establishment of rapport is important as well. Also one needs to use simple verbal forms, easily understandable to a child of a particular linguistic, social, and cognitive level. The interviewer needs to test the child’s ability to distinguish between true and false statements, considering that the ability to understand what a lie is still developing at the age four and five years. These aspects of the interview are described in more detail in Mazzoni (2001; see also Bull, 1995), in which the good interviewer is described as one who can facilitate the narrative, not one who guides and leads it.

Good interviews are rare, and this has brought about the problem of the reliability of child eyewitness testimony. Usually verbal narratives of children are relatively short (Goodman & Reed, 1986), but also quite accurate (Dent, 1988; Cole e Loftus, 1987). Repetitions of the same report do not present new information or intrusions. The child’s memory of an event improves if the child was in the event (Fivush, 1993) and if the child is active in the episode, instead of just being a witness (Tobey & Goodman, 1992). The best way to obtain an accurate memory is to let the child recall the episode freely, whereas accuracy decreases substantially when children are asked specific questions or try to recognize faces (Brainerd, Ornstein, 1991; Dent, 1992; Goodman, Bottoms,

Swartz-Kennet, Rudy, 1991; Naylor, 1989). However, in freely recalling scenes children tend to leave out peripheral information, elements that are not central or salient, and this limitation can require using more specific questions under some circumstances (Goodman et al, 1987).

One major problem is the fact that young children tend to accept wrong information that is suggested by adults (Ceci, Leichtman, 1992; Goodman, Reed, 1986). However it also seems that when the degree of involvement in an event is particularly strong, even very young children show a high degree of resistance to adult suggestion (Brigham et al., 1986; Rudy e Goodman, 1991; Goodman, Rudy, Bottoms e Aman, 1990). Several studies have examined the role the type of information and of how misleading information is presented on children suggestibility. Two major variables are the credibility of the person that gives the suggestion (Ceci e Bruck, 1993; Belli, Lindsay, Gales e McCarthy, 1994) and the salience of the information. Other studies (e.g. Mazzoni, 1995) have shown that the mechanisms that make children suggestible change with age: at age six misleading information replaces parts of the memory of the original event; at age ten it just gets added to the memory of the original event. Memory suggestibility is overall a very important characteristic of the child and needs to be taken into serious consideration when interviewing children.

The validation of the report

Lack of accuracy of the eyewitness report is an ever-present problem in court cases, and in particular when child sexual abuse cases are considered, since often the child is the victim and also the only eyewitness of the facts. It is very common to come up with wrong decisions about child sexual abuse cases. For this reason the German courts have been using for many years a specific technique for the validation of eyewitness statements. It is based on a content analysis of the statements (Statement Validity Analysis, or SVA) that can reduce the probability of judgmental errors. The SVA is a structured method for collecting and analyzing the information provided by a witness and is often used in sexual abuse cases. It evaluates the validity of the report, not the general credibility of the witness. Thousands of interviews have been conducted and analyzed using such method, which is currently quite popular.

It is misleading to think that a witness has to be coherent in order to be credible: only that which was reported during the interrogation needs to be coherent (Ghetti & Agnoli, 1998). One should then only talk about the credibility of the report and not about the credibility of the person. That is why only the content of the report is analyzed using the SVA

The SVA includes the following five phases: a) careful assessment of the case; b) quasi-structured interview of the child; c) analysis of the content of the interview based on a number of specific criteria; d) validity assessment; e) integration of c) and d). Information about the case is gathered using all possible sources: police reports, files of child protection agencies, psychosocial anamnesis of the child and other people involved in the case, information provided by the school and by the court, etc. (Steller & Boychuck, 1992). All of this information will allow the formulation of various alternative hypothesis of what happened. The interview is used to evaluate and compare the various alternative hypotheses.

The format of the interview was designed to obtain as much information as possible through free narration. Questions are structured in a “funnel,” starting from the

more general and moving to the more specific (Steler & Boychuck, 1992). The content of the report is then carefully analyzed using a method called Criteria Based Content Analysis (CBCA). The method is based on the hypothesis proposed by Undeutsch (1989), that reports of real experiences are qualitatively different from reports of invented experiences that have been made up by the child or by adults. The content is then evaluated according to nineteen criteria, grouped into 5 categories. These are intended to allow discrimination between true and false statements (Steller & Koehnken, 1989). A validity checklist has been introduced to add information to the results of the CBCA. It represents a way to guarantee the accuracy of the evaluation of the case by requiring that the decision be postponed until all aspects of the SVA are carefully considered. The SVA and the CBCA (which represents the core part of the method) are valuable aids in that they systematize and organize the interview and the evaluation procedure and control misleading variables such as personal bias and prejudice.

The CBCA

The CBCA allows the examiner to evaluate the presence of 19 criteria. The higher the level of the criteria in the statement, the higher the likelihood that the statement is credible. Five categories organize the 19 criteria. The first concerns the general characteristics of the report and requires a careful analysis of the whole report. It includes the following criteria: logical structure, unstructured production, and presence of details. The second and third categories refer to specific aspects of the deposition. The fourth category includes motivational aspects and the fifth includes elements that specifically refer to the offense. The criteria and the five categories are reported in more detail in Table 1.

The first criterion concerns the coherence and internal consistency of the statements. The second criterion indicates a statement that is not structured as a list of events. The third criterion is that the statement includes specific descriptions of the location, time, objects, people etc.

Criteria 4-7 belong to the second category. Criterion 4 refers to the connection (spatial and temporal) of the critical event with more typical events of the child's life. Criterion 5 refers to physical interactions among people. Criterion 6 refers to the presence of conversations in the statement. Criterion 7 is met if the report contains unexpected complications that usually are not present in false, fabricated statements.

The third category includes criteria 8-13. Criterion 8 refers to the presence in the statement of details that are not inferable from the mental schema of the event, or from the context. Superfluous details are considered in the criterion 9: these details are consistent with the context, but are irrelevant to the main event. The tenth criterion is met when the child reports facts or words that he or she does not understand (for example, misinterpreting a sexual intercourse). Criterion 11 refers to a situation in which the child reports conversations or events of a sexual nature that had not happened during the event described in the interview. Criteria 12 and 13 refer to statements that include comments on the mental condition of the people involved in the episode. Criterion 12 is met if the child talks about personal emotions and feelings. Criterion 13 refers to emotions and feeling of the alleged perpetrator.

The fourth category includes elements that refer to the child awareness of the content his/her narrative and behavior. They include elements like spontaneous corrections, comments about memory failures or lack of memory, comments about the

inadequacy of his/her behavior during the episode, and positive comments about the alleged perpetrator. The fifth category includes only one criterion (19) that refers to the description of the abuse in a detailed way. It is rarely present in credible reports, though.

Usually two judges evaluate the presence of the 19 criteria in the report of a witness. A score of 0 is assigned to a criterion if it is absent; a score of 1 is assigned if it is present, a score of 2 is assigned if it is strongly present. Despite the apparent simplicity of the scale, deciding about the presence of the 19 criteria is very complex, because there are no specific rules that can clearly indicate when a criterion is met. In addition, the final evaluation of the credibility (true/false) of the report is complicated by the lack of decisional criteria, which are left to the subjective judgment. For example, some authors (e.g. Esplin et al, 1988) have suggested that the decision can be based on the sum of the scores obtained on the first five criteria. Others (eg. Yuille, 1988) suggested that in order to make the decision one should consider the first five criteria plus 2 others at random. Currently, there are also no formal rules for combining the scores of the 19 criteria. Future studies should address the issue of the importance and weighing of the single criteria. And there is a great need for studies that test the validity and reliability of this instrument that is extensively used in the courts.

The purpose of the present study was to assess the validity of the CBCA by testing whether it can discriminate between narratives of false and true events in 7-year-old children. In a previous study, Mazzoni and Pezzati (in press) demonstrated that only few criteria of the CBCA can discriminate between narratives of false and true events in 4-year-old children. In the present study we examined whether the same criteria were also capable of discriminating between true and false narratives in older (7-year-old) children, and whether at this age additional criteria could discriminate between narratives.

The narratives were about events that were as much as possible traumatic events, in order to make them similar to a sexual abuse episode (e.g. drawing blood, being beaten by a peer, being bitten by a dog, going to the dentist, going through surgery, etc). Three independent judges examined the narratives (see also Steller, Wellerhaus & Wolfe, 1988). Children were asked to tell a story about an event they had really experienced and a false event they were asked to invent.

The true event had all happened during the previous 12 months and children were asked to tell what happened starting with a completely free recall. The narration was prompted by saying "And then?" "What else happened?" "Is there anything else you want to tell me?". After one week, the children were asked to invent an event. Then each narrative was analyzed using the criteria of the CBCA

Method

Participants

Forty-five children of two schools of the Cosenza province took part in the study. Fifteen children were not able to tell an invented story or to understand what the task was. Thus, the analyses were done on 30 children. Each of the children narrated a true event and invented a false event.

Material

Children received a questionnaire asking which of the following event had happened to them during the previous 12 months:

Cut one's finger with a knife

Being bitten by a dog
Falling from the bicycle
Swallow poison
Fall from the stairs
Getting an electric shock
Being stung by an insect
Being burnt by a match
Breaking an arm or a leg
Being beaten by another kid

For each event, children had to state when it happened to them and whether they were scared. Depending on the responses, six events were selected: falling from the stairs, being stung by an insect, falling from the bicycle, getting an electric shock, being burnt by a match. These events were selected because half of the children stated that they experienced at least one of them, and half of the children stated that they did not experience at least one of them. Thus, the number of times the event was true was equal to the number of times the event was false. This allowed us to collect an equal number of true and false narratives.

Procedure

The experiment comprised three phases held during a one-month period.

Phase 1. All children attending the two schools were presented with the questionnaire. Children were told that we were studying how children narrate a story. The questionnaire was presented in class, since all children could read and write easily. When children indicated that an event had happened to them during the previous 12 months, they were asked to specify when and to indicate whether it had scared them. The truthfulness of the answers was checked with parents and teachers. The six episodes were selected on the basis of the frequency with which children stated the various episodes had happened or had not happened to them. Selection was made to meet the criterion of having the same number of times in which each event was experienced and not experienced. Phase 1 lasted one week.

Phase 2. In this phase, children were asked to narrate the true event. Interviews were done individually in a classroom. Each interview started with the question: "You told me that you fell down the stairs and that you hurt yourself. Tell me more about what you remember of this episode, so that I can better understand what happened". Each narrative was recorded. If the event occurred more than once during the 12 months, children were asked to tell the one that they remembered best.

Phase 3. One week later, the children were asked to narrate the event that they did not experience, as if it actually occurred. The one-week interval ensured enough time so that children were not influenced by the narrative style they used when talking about the true event. The interview started by asking "You told me that ... (the event) did not happen to you. Now imagine that it actually happened to you. Try to tell me what happened, in the same way as when your teacher asks you to tell a story about some event."

During the interview about both true and false events, when the child was stopping without telling too much, the experimenter asked: "And then...?". Depending on the answer, the experimenter then asked more specific questions about the event (e.g.

“Where were you?”). All interviews were preceded by a period of time devoted to establishing rapport between the child and the experimenter.

Results

Scores: All narratives were tape recorded and then transcribed. A score was assigned for each of the 19 criteria of the CBCA, and these scores were then summed to get a global score for each narrative. A score of 0 was assigned when the criterion was absent; a score of 1 was assigned when the criterion was present, and a score of 2 when the criterion was strongly present. The maximum score for each narrative is 38. The minimum is 0.

A Spearman correlation was computed between the scores assigned by the two judges to the true narratives, and another one was computed between the scores assigned to the false narratives. Data are reported in Table 2.

Insert Table 2 here

Given the correlation, the scores for true and false narratives assigned by the two judges were averaged. Data are reported in Table 3.

Insert Table 3 here

The following analyses examine which of the 19 criteria are responsible for the difference. A 19 (criteria) x 2 (type of narrative) within-subjects ANOVA showed that the two narratives received significantly different scores, $F(1,57) = 25.49$, showing that the criteria are overall more present in true narratives. There was a significant effect of criteria, $F(18,1026) = 98.33$, $p < .001$. This suggests that some criteria are more present than others. The interaction was also significant, indicating that true and false narratives receive different scores in some criteria but not in others. In order to understand which criteria discriminate between true and false statements, and which do not, we computed 19 simple effect tests, which are reported in Table 4 along with the mean scores of the 19 criteria.

Table 4 shows that six criteria discriminate between true and false narratives: n. 1; n.2; n. 3; n.12; n. 13 and n. 18.

Discussion

The growing number of child sexual abuse cases in which the only witness is also the victim increases the need to develop and test methods that can discriminate between true and false reports and statements by children. In the present study we used the CBCA, a method to analyze the content of a narrative. We assessed whether the CBCA can discriminate between narratives of events that are experienced (true) and narratives of events that are not experienced (false) in 7-year-old children.

Previous studies on the use of CBCA are sparse and do not provide conclusive data. These studies examined children older than 7 (Koehnken, Schimossek, Ashermann e Hofer, 1995; Ruby e Brigham, 1998), or have included samples of children of several different ages, between 3 and 12 years old (Boychuck, 1991) or between 3 and 15 years old (Esplin et al, 1988) (see also Ghetti e Agnoli, 2000). Other studies have tested the CBCA in a more homogenous sample for age (six to nine years, Yuille, 1988). However none of these studies allow the conclusion that the CBCA is valid at a specific age (e.g. age of 7).

The current study, conducted on a homogenous sample of 7-year-old children, showed that children at this age tell a true story in a way that is appreciably different from the way they tell a false story. True and false narratives could be discriminated both

on the basis of the global score assigned by the judges, and on the basis of the scores of six criteria, n.1, n. 2, n.3, n. 12, n. 13 and n. 18. These criteria refer to the general characteristics of the narrative (Crit. 1: logic structure of the narrative; Crit. 2: nonstructured narrative; Crit. 3: number of details); the specific content of the narrative (Crit. 12: description of the mental state of the child; Crit. 13: attributing a mental state to the alleged perpetrator), and cognitive/moral comments (Crit. 18: forgiving the perpetrator). Thus, true stories are structured with a stronger internal logic (1), that makes the narrative more coherent, better structured, and better interconnected. In addition the narrative structure is not simple, but relatively complex (Crit. 2). The structure does not follow a temporal linear progression of events (e.g. and then... and then... and then...). Instead, the connection is based on a mixture of causal as well as temporal links. The temporal links can go forward but also backward, creating loops and parenthetical comments. Finally, the structure is enriched by details. Children talk about the location, about the people, objects, weather more easily when they tell a true story than when they have to invent a story about an event that they did not experience.

In a previous study Mazzoni & Pezzati (in press) found that even at the age of 5, children create different structures for true and false narratives. At that age as well, true stories received significantly higher scores in Crit. 1 (logic structure), Crit. 2 (complex structure) and Crit. 3 (number of details). The same three criteria discriminate between true and false narratives also among older children. Being able to create a coherent and interconnected narrative with a relatively complex structure is a cognitive task that children handle better when they tell true stories. It is more difficult when telling stories about events that have not been experienced.

A further element that discriminates true from false narratives at the age of 5 is n. 12 (talking about personal feelings and mental states). This criterion is more present in true than in false narratives. It is probably easier to remember (and then talk about) emotions (e.g. "I was scared."; "I was angry...") that are actually experienced than inventing or inferring emotions within a given situation that was not experienced.

Finally, at age 7, but not at age 5, it becomes possible to start expressing emotions or mental states of other people that have been part of the experience, whereas it is still very difficult to invent or infer mental states in others in a situation that has not been experienced (Crit. 13). Children also start making comments about the degree of guilt of the perpetrator (Crit. 18). However both criteria are still relatively infrequent even in true narrative, and the significant difference obtained between true and false narratives depends on the fact that they are completely absent from false narratives. This does not diminish the fact that at this age children are able to comment about their own emotional and mental states, they start being sensitive to the thoughts and emotions of the perpetrator, and to the perpetrator's guilt.

In conclusion, these data confirm previous results (Mazzoni & Pezzati, in press) and indicate that one can rely on the first three criteria of the CBCA to discriminate true events from false events. The three additional criteria that appear able to discriminate at age 7 suggest that at different ages probably different additional criteria can be used. If these results are replicated, it can be concluded that only a few, basic, criteria are really crucial in making the discrimination between true and false narratives. The ones that are added at various ages might simply be due to an improvement in the ability to use

language and to an increasing awareness of mental states and other elements (e.g. moral judgment).

This hypothesis is supported by data from previous studies showing that Logic Structure can discriminate between true and false statements in children 6 to 12 (Yuille, 1988; Steller et al, 1988), and in children that had really experienced sexual abuse (Esplin et al, 1988; Boychuck, 1991; Lamers-Winkelman e Buffin, 1996). This criterion has high reliability (Anson et al, 1993). Similar to Crit. 1, Crit. 3 was found to discriminate in almost all the studies done until now on CBCA. Other data (Lamers-Winkelman e Buffin, 1996) support the importance of Crit. 3, showing that this criterion is found in true stories independently of the age of the witness. That study showed that criterion n. 12 is present independently of the age of the witness. This criterion was found to discriminate not only in our previous study (Mazzoni & pezati, in press), but also in the Steller et al (1988), Esplin et al (1988) and (Lamers-Winkelman & Buffin, 1996) studies.

The results of the current study are encouraging and in line with previous data. However it is important to emphasize the dangers of an incorrect use of these data. Today, in Italy and elsewhere, there is disagreement on how to score the CBCA and on how to use the results obtained in the experimental research. Some have suggested that using a 0/1/2 scoring system is not appropriate when using the C.B.C.A. in court cases (Steller, personal communication). According to Steller, the CBCA cannot be considered a scale, because it is a clinical tool that is not psychometrically tested and that should be used only as an aid in deciding about the truth of a statement.

The lack of psychometric testing creates problems when trying to use the global score obtained with the CBCA as a discriminative tool. However, it does not hinder the possibility of comparing the presence of each criterion in true and false narratives, since the weight of the criterion in this case remains the same. Steller (Steller et al, 1988) used this methodology in his studies, reaching conclusions similar to ours.

The critiques to the CBCA have led to the development of other scales with better psychometric properties (Spoerer, personal communication, 2001). These scales are still under development. Meanwhile, the popularity of the CBCA increases both in Italy and abroad.

In Italy, the CBCA is used mostly by people who are not experienced or appropriately trained in its use. There are no professionals trained in the use of the CBCA in this country. Those who use it, seem to underestimate the danger of using in real life and in real court cases (as they often do, see Pezzati, 2002) an instrument that they do not know and that they have not been trained to use. It is one thing is to use the CBCA in an experimental context for research, another thing is to use it to decide about the fate of a number of people. Although research is necessary to understand the limits and potentials of the CBCA, the results cannot be immediately transposed and used in court cases.