The purpose of this work is, on the one hand, to describe Criteria-Based Content Analysis (CBCA), a procedure focusing originally on the assessment of children’s testimony credibility, and on the other, to present some of the research on this topic. In recent years, however, there has been increasing interest in the application of this procedure to adults, so that another objective of the study was to discuss studies on the use of this statement credibility analysis technique with adults. Finally, we discuss some disadvantages of the technique and future lines of research in relation to its use.

Este trabajo está dirigido por un lado a la descripción del Análisis de Contenido Basado en Criterios (CBCA), procedimiento enfocado originalmente a la evaluación de la credibilidad del testimonio de niños, y por otro, a la presentación de algunas investigaciones realizadas en torno a este tema. En los últimos años, sin embargo, ha habido un creciente interés por generalizar la aplicación de este procedimiento a adultos, por lo que otro de nuestros objetivos será comentar los estudios encaminados a emplear esta prueba en la evaluación de la credibilidad de las declaraciones de adultos. Finalmente comentaremos algunos inconvenientes de esta técnica y futuras líneas de investigación.

It is not uncommon for police investigations to have access to no more than the testimonies of the victim and the accused as evidence in a crime. Given this situation, several researchers have worked on the development of systematic methods that help to distinguish honest testimonies from those that have been fabricated. Vrij (2000), and Vrij, Edward and Bull (2001) have classified these procedures in three groups. The first of these includes procedures for recording and analyzing the psychophysiological activity of the person who is lying; the second group concerns the examination of the witness’s non-verbal behaviour (Vrij, Edward, Roberts & Bull, 2000); and the third, on which we shall concentrate here, refers to the study of the content of the witness’s statement (Masip, Sperer, Garrido & Herrero, 2005; Ruby & Brigham, 1997).

DEVELOPMENT OF CRITERIA-BASED CONTENT ANALYSIS (CBCA)

Statement Validity Assessment (SVA) is the technique most widely employed for assessing the truthfulness of verbal statements (Vrij, 2000). The SVA was developed in Germany, based on the clinical experience of several psychologists. It was around 1950 that Udo Undeutsch first described SVA (see Undeutsch, 1989), which was subsequently modified, on the way to its current form, by Steller and Köhnken (1989) and Raskin and Esplin (1991). SVA was initially developed for assessing the verbal statements of minors who had been the victims of sexual abuse. However, in recent years it has been attempted to validate and generalize the application of this instrument for adults (Vrij et al., 2001; Vrij et al., 2000). Despite the fact that it is an instrument widely used in the forensic context as psychological evidence, it should be considered not as a test or standardized scale, but as semi-standardized method for assessing the credibility of statements (Steller, 1989). The development of SVA is based on what Steller (1989) has called the Undeutsch hypothesis. According to this hypothesis, a testimony based on a real experience differs in quality and content from a testimony based on an imagined event.

Criteria-Based Content Analysis (CBCA) is the core component of SVA, and this has led to its being the element most frequently studied by researchers (Ruby & Brigham, 1997); it also serves as the focus of the present work. SVA is made up of three mutually dependent components: a) a structured interview with the victim, b) CBCA, which assesses the content of the person’s testimony, and c) the integration of CBCA with the information obtained through a set of questions called the Validity
Checklist, so that the information provided by content analysis of the statement is combined with other relevant information of the case and with information obtained from the exploration of the interview or interviews previously carried out (Horowitz, 1991).

The interview should precede application of the CBCA criteria. The basic objective is to obtain material on which to apply these criteria. For the interview to be appropriately conducted, it is important for the interviewer to be familiar with the content of the criteria. Likewise, the interviewer should try to obtain the greatest possible amount of data by using an interview designed to maximize the quantity of information provided by the witness and minimize any type of contamination generated by either the interviewer him/herself or any other adult (Raskin & Esplin, 1991).

CBCA is applied to the content of the testimony, and its purpose is to determine whether the quality and specific content of that testimony are indicative of an account generated from memory traces or of one that is the product of invention, fantasy or the influence of another person. Any analysis made using CBCA will be influenced by the interview characteristics and by what the subject has or has not experienced. Throughout this process, it is important that the interviewer takes into account the person’s age, experience and cognitive ability level (Raskin & Esplin, 1991). One of the main limitations of CBCA concerns the fact that it is applied to situations in which witnesses have information on the basis of which they can invent a charge that incorporates some of the criteria. For example, a child who has been sexually abused on a previous occasion may provide a false testimony that appears convincing as it is based on the memory traces from that previous experience. This aspect should be borne in mind on reviewing the case by means of the Validity Checklist (Raskin & Esplin, 1991). The verbal content of the statement is analyzed through the application of a series of 19 criteria (see Table 1), organized in five broad categories, and with the purpose of differentiating between true statements and fabricated statements. The basic idea is that a true testimony contains a greater number of criteria (for a detailed description of these criteria, see Steller and Köhnken, 1989).

Analysis of the interview by means of the 19 CBCA criteria is carried out by giving numerical scores to each one of the criteria. Steller’s (1989) proposal is to assign 2, 1 or 0 points depending on whether the criterion is strongly present, present or absent in the statement, while other authors (Lamb, Sternberg, Esplin, Hershkowitz, Orbach & Hovav, 1997) propose awarding 1 for present and 0 for absent. Neither the interview nor the results obtained from the criteria are completely valid until they have been put into context by means of the Validity Checklist, which is made up of four general categories of information (Steller & Köhnken, 1989):

a) Psychological characteristics. In this category it is important to assess the appropriateness of language and affect and susceptibility to suggestion.

b) Interview characteristics. The assessor should analyze the quality of the interview, rating the type of questions asked (suggestive, leading or coactive questions) and its overall appropriateness.

c) Motivation for making false accusations. The information in this category should help to rule out those aspects of a motivational nature that may be influencing the person to provide a false testimony. It

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>CRITERIA-BASED CONTENT ANALYSIS (CBCA), MODIFIED FROM STELLER AND KÖHNKEN (1989)</th>
</tr>
</thead>
</table>
| General Characteristics | 1. Logical structure  
2. Unstructured production  
3. Quantity of details |
| Specific Content | 4. Contextual embedding  
5. Description of interactions  
6. Reproduction of conversation  
7. Unexpected complications during the incident |
| Peculiarities of Content | 8. Unusual details  
9. Superfluous details  
10. Accurately reported details misunderstood  
11. Related external associations  
12. Allusions to subjective mental state  
13. Attribution of the accused’s mental state |
| Motivation-Related Content | 14. Spontaneous corrections  
15. Admitting lack of memory  
16. Raising doubts about one’s own testimony  
17. Self-Deprecation  
18. Pardoning the accused |
| Specific Elements of the Offence | 19. Specific details of the offence |
should also be borne in mind that a minor can be under pressure from a third person to make a false statement. An important aspect of this category is the assessment of the context in which the statement is generated.

d) Aspects related to the investigation. This section is designed with the aim of rating the consistency between previous statements and investigation results and medical reports.

The assessor must analyze the information from these four categories and, on the basis of this analysis, determine whether this information supports the person’s testimony. Likewise, the objective of the Validity Checklist is to rate various explanatory hypotheses, examining all the available information on the case. Raskin and Esplin (1991) propose five hypotheses that should be considered by the assessor:

a) The statement is valid, but the child has substituted the identity of the aggressor by that of a different person.

b) The statement is valid, but the child has been influenced or has invented additional information that is not true.

c) The child has been put under pressure by a third person to formulate a false version of events.

d) Due to personal interests or to help third persons, the child has given a false statement.

e) As a consequence of psychological problems, the child has given a testimony based on fantasy or invention.

It is important to stress that the purpose of SVA is to assess the credibility of the content of statements, and not to assess the credibility of the persons themselves (Steller & Köhnken, 1989).

Another serious limitation of CBCA is the lack up to now of a decision rule that helps to establish how many criteria determine whether a statement should be classified as credible or not credible; even less consideration has been given to the weight each criterion should be assigned. Alonso-Quecuty (1999) proposes that the weight of each criterion be assigned on the basis of diverse factors, such as: number of previous interviews the child has had, complexity of the incident, age of the minor, and time elapsed since the event. Once the CBCA criteria and the Validity Checklist have been applied, the final result of the analysis permits the statement to be qualitatively classified according to five categories (Alonso-Quecuty, 1999; Steller, 1989):

- Credible.
- Probably credible.
- Indeterminate.
- Probably not credible.
- Not credible.

RESEARCH IN RELATION TO CBCA

Recent years have seen an increase in the number of studies on CBCA due to its extensive use in the forensic context. Studies have been of two basic types: 1) those that use real cases of minors allegedly the victims of sexual abuse, and in which other elements of the case are used as measures of truthfulness; and 2) experimental studies in which subjects are induced to manipulate their statement, providing either a true or false testimony (Ruby & Brigham, 1997).

Studies with children

Given the fact that CBCA was designed to be applied to the statements of minors alleged to be the victims of sexual abuse, the majority of published studies have employed samples with these characteristics. Let us briefly consider some of these studies carried out with children.

An important study, insofar as it focused on children presumed to have suffered sexual abuse, was that of Lamb et al. (1997). Their sample was made up of 98 Israeli children (28 boys and 70 girls) aged between 4 and 13 (mean 8.72). As measures of the truthfulness of the statement they employed other elements of the case, including material or physical evidence, the accused’s testimony, and so on. As predicted, there was greater presence of the CBCA criteria in the credible accounts (mean 6.74) than in the not-credible accounts (mean 4.85). Nevertheless, the authors state that the differences found were not as significant as those of previous studies.

More recently, Santtila, Roppola, Runtti and Niem (2000) analyzed the effect of age, verbal ability (measured with the WISC-R verbal scale) and interviewer’s emotional style on the presence of CBCA criteria in the statements of 68 children from three different age groups: 7-8, 10-11 and 13-14 years. In their experiment, they asked each child to give an account of two personal experiences, one real and one false. The results showed a correct classification rate of 66%. They also found that age and verbal ability of the minor, as well as interviewer behaviour, influenced the number of CBCA criteria present in the statements, regardless of whether
these were true or false. Likewise, they found that different criteria appeared in the statements depending on participants’ age range. These authors suggest integrating the information from CBCA with that obtained through the Validity Checklist, and conclude that judicial decisions should not be based exclusively on the results provided by CBCA-SVA.

Research has also focused on analyzing the influence of the familiarity of the event to be recalled on the presence of CBCA criteria. For example, Pezdek et al. (2004) carried out an experiment with 114 children, hypothesizing that descriptions of familiar events were more likely to be classified as true than descriptions of unfamiliar events. The results suggested a greater presence of CBCA criteria in accounts of familiar events than in accounts of unfamiliar events.

In a similar line, Blandon-Gitlin, Pezdek, Rogers and Brodie (2005), with a sample of 94 children, analyzed the interaction between familiarity of the event and its veracity. They found the scores obtained through CBCA to be more strongly influenced by the familiarity of the event than by its truthfulness. In both this study and the previous one, the authors concluded by suggesting that CBCA, in its current form, is of limited utility as a tool for assessing statement credibility in minors.

Studies with adults

Given the good results obtained with children, it has been attempted to apply it to adults also. However, there are fewer studies than in the case of minors. Let us consider a few of them.

Some research has concentrated on identifying the criteria most commonly found in true statements, which would consequently be the most sensitive ones in the discrimination between true and false testimonies. For example, in a meta-analysis, Ruby and Brigham (1997) found the criteria most frequently found in true statements to be 1, 2, 3, 4, 5, 6, 7, 8, 12, 14 and 15. On the other hand, Köhnken, Schimossek, Aschermann and Höfer (1995) found that true testimonies included a larger number of details (Criterion 3) and were unstructured (Criterion 4), and that people had a greater tendency to admit lack of memory about the event (Criterion 15). They also found five of the six additional criteria included in their experiment to be significant, though while four of them (expression of insecurity, reporting style, justification of memory lapse and clichés) were significantly more likely in false statements, the fifth (repetitions), contrary to predictions, was significantly more frequently found in truthful accounts (for a detailed description of the additional criteria included in this study, see Köhnken et al., 1995). Likewise, Vrij, Akehurst, Soukara and Bull (2004a) reported that Criteria 1, 3, 4, 5 and 6 were the most effective for differentiating between true and false testimonies.

The vast majority of studies that analyze the effectiveness of CBCA have used samples of Europeans or other Caucasian people. Therefore, Ruby and Brigham (1998) decided to explore the differences that may arise between subjects from different ethnic groups. These authors started out from the notion that there are differences at a verbal level, in terms of style and content, between accounts provided by persons of different racial origin. The hypothesis they proposed was that, since CBCA was developed in a white European culture, its application would only be effective for discriminating between the testimonies of this type of subject, and that accounts by black subjects would include significantly fewer criteria. The results showed that CBCA functioned differently according to a person’s race, and that different criteria were better predictors of truth for one ethnic group than for the other. In the true accounts given by the black subjects, the criteria that most frequently appeared were 3, 6, 12, 14, and 17, though a comparison of true accounts by white subjects with those from black subjects showed no criterion that appeared significantly more frequently. Considering the testimonies of the subjects of both races jointly, it was found that certain criteria appeared significantly more in true statements (Criteria 2, 5, 7, 8, 9, 14 and 15). However, in false accounts there was also greater presence of certain criteria: 1, 4, 11, 12 and 17. As regards the classification of the statements, the results were not conclusive. The researchers found that if they took as true all those statements in which 5 criteria were present, the percentage of correct classification for true statements was 89%; however, with this rule, a high percentage of false accounts were also classified as true (92%). On employing stricter decision criteria (taking into account the presence of 6 or 7 criteria), the number of true accounts correctly classified fell, and the number of false accounts correctly classified increased - that is, with this criterion fewer false accounts were classified as true.

Research has also revealed that there are differences in CBCA scores on comparing the results for children and adults. This is the case, for example, of the study by Vrij et al. (2004a). It has also been found that previous
knowledge of the content of the CBCA criteria has a negative influence on the validity of the instrument, and that subjects instructed to lie and who know the criteria in advance can provide statements that appear to be true (Vrij, Akehurst, Soukara & Bull, 2002).

The effectiveness of CBCA has also been compared with that of other procedures for assessing the content of statements, and researchers have explored whether the combination of these methods improves the classification of testimonies. Specifically, reality monitoring (Johnson & Raye, 1981) is the procedure with which it has most commonly been compared. Reality monitoring developed within the basic research context, and was initially applied in the clinical field (where it stimulated extensive research, e.g., Bentall, Baker & Havers, 1991; Brebion, Smith, Gorman & Amador, 1997; Harvey, 1985; Johnson, Raye, Hascher & Chromia, 1979; Raye & Johnson, 1980), before being employed in the forensic context (see, for a review, Mitchell & Johnson, 2000). It postulates that memories of what was actually seen have different characteristics from internally-generated “memories”. The proposal of reality monitoring is that memories of external origin will have more contextual and sensory attributes, greater semantic detail, and less information on cognitive operations than memories of internal origin (Johnson & Raye, 1981).

One study that compares the results of CBCA and those of reality monitoring is that carried out by Sporer (1997). In his experiment, Sporer used a sample of 40 psychology students (20 men and 20 women), and participants were instructed to recount two personal experiences: one true and the other false. The author’s objective was to explore the efficacy of CBCA and reality monitoring in the discrimination of fabricated and truthful accounts and to check whether the combined use of the two instruments improved the classification of the accounts. The results showed that CBCA was effective in 65% of total classifications, with 70% effectiveness in the classification of the true accounts and 60% effectiveness in that of the false accounts. As regards reality monitoring, 71.3% of the statements were correctly classified. Of the true statements, 75% were correctly classified, and of the false accounts, 67.5%. On combining CBCA and reality monitoring, the percentage of correct classification rose to 79%. In a later study, Vrij, Akehurst, Soukara and Bull (2004b) found that the true testimonies obtained higher scores both in CBCA and the criteria of reality monitoring, with classification effectiveness of 60% and 74%, respectively. Nevertheless, on integrating the results of the two instruments, no improvements were found, and the percentage of classification remained at 74%, so that the combination of the two techniques did not produce improvements on this occasion.

Various studies have also considered the option of combining procedures for assessing the verbal content of statements with behavioural indicators of deception. Vrij et al. (2001) found that those who lied scored lower in the CBCA and reality monitoring criteria and were more likely to present certain behaviours indicative of deception, such as waiting a long time before answering, talking more quickly, and so on. They even found that both CBCA and reality monitoring were the most sensitive tools in the detection of deceit in relation to other variables, such as non-verbal behaviour. In a previous study, Vrij et al. (2000) found that the joint use of verbal and non-verbal indicators of deception (CBCA and reality monitoring) led to an increase in the percentage of correct classifications. These results were ratified in a more recent study by Vrij et al. (2004a) that examined the verbal and non-verbal behaviour of children and adults.

CONCLUSIONS

Criteria-Based Content Analysis (CBCA) is still far from being a totally effective tool in the detection of deceitful testimonies, and requires a good deal of refining. Regardless of whether this technique is applied to children or adults, there are many factors that exert a negative influence and can affect its results. As shown by research, there are individual differences, such as age, verbal ability, interviewer’s attitude (Santilla et al., 2000), familiarity of the event (Blandon-Gitlin et al., 2005; Pezdek et al., 2004), previous knowledge of the instrument, (Vrij et al., 2002) and ethnic group of the person (Ruby & Brigham, 1998), that have to be taken into account and controlled as far as possible when this technique is employed, and which, therefore prevent the immediate individual application of CBCA.

However, and although studies show that true statements contain a larger number of criteria than fabricated statements, the major disadvantage of CBCA is that there is no general consensus establishing a minimum number of criteria a statement should include in order to be classified as credible, or the weight that should be attributed to each of them. Landry and Brigham (1992) have proposed as a minimum the presence of five criteria for a statement to be classified as truthful. However, other au-
thors manipulated in their experiments the number of criteria in the classification of statements, and their results were not as conclusive as expected (Ruby & Brigham, 1998). Likewise, there is still a need to define the number of criteria that should be included in assessment of the testimonies of adults. Given that CBCA was developed for assessing the statements of minors, it is likely that some criteria do not work with adults, as is the case of Criterion 10 (Accurately reported details misunderstood). Thus, there is also a need for studies aimed at the definition of a group of criteria applicable to the testimonies of adults. As regards the weight each criterion should receive, in this aspect research is even further away from its objective, that of setting a general parameter. Nevertheless, certain criteria have been found to discriminate better than others between true and false testimonies (Ruby & Brigham, 1997, 1998). It may be that some of these criteria should receive higher scores in the general assessment of CBCA; however, it is essential to carry out further research with a view to clarifying this issue.

In accordance with the situation as presented here, a viable alternative in the detection of false testimonies is the combination of various techniques, such as reality monitoring and behavioural indicators of deception (Vrij et al., 2001; Vrij et al., 2000), which, as we have seen, on the majority of occasions improves the classification of statements.

Given the disadvantages mentioned, what does clearly emerge is that CBCA should be considered exclusively as a support tool, and never as the sole instrument on which to base judicial decisions (Santtila et al., 2000), at least for the present time.

REFERENCES


