

Is This Testimony Truthful, Fabricated, or Based on False Memory?

Credibility Assessment 25 Years After Steller and Köhnken (1989)

Renate Volbert and Max Steller

Institute of Forensic Psychiatry, Charité, Berlin, Germany

Abstract. In 1989, Steller and Köhnken presented a systematic compilation of content characteristics for distinguishing between truthful and fabricated testimonies (criteria-based content analysis – CBCA) designed to be applied within a more comprehensive overall diagnostic procedure known as statement validity assessment (SVA; Steller, 1989). The subsequent 25 years have seen a marked increase in knowledge about the distinction between experience-based and non-experience-based statements. This supports the SVA approach and permits a better explanation of the underlying processes. The rationale of CBCA is that a true statement differs in content quality from a fabricated account because (a) a truth teller can draw on an episodic autobiographical representation containing a multitude of details, whereas a liar has to relate to scripts containing only general details of an event; and (b) a liar is busier with strategic self-presentation than a truth teller. The present article proposes a modified model of content characteristics that pays greater attention than before to these underlying processes. SVA takes into account that content quality is influenced not only by the veracity of a statement but also by other (personal and contextual) variables that need to be considered in the individual credibility assessment. Theoretical analyses and empirical research do not indicate comparable qualitative differences between true statements and those based on false memories. Witnesses giving testimony based on false memories do not fabricate false statements actively, and they make no effort to conceal a deception; they are not deceiving but mistaken. In these cases, a noncritical application of content criteria can lead to false results. To examine the hypothesis that a statement is based on a false memory, it is necessary to focus on the way in which the statement has emerged and evolved.

Keywords: criteria-based content analysis, statement validity analysis, credibility assessment, false memory

Assessing the veracity of testimony becomes crucial when one statement contradicts another during criminal proceedings. This situation occurs frequently in sexual offenses: The accused denies the offense, and – due to a lack of material evidence and eye witnesses – the only evidence is the incriminating testimony of the victim. For several decades now, courts in Germany and in several other European countries have addressed this issue by appointing psycho-

logical experts to deliver expert opinions on the credibility of the testimony given in a particular case.

In 1989, Steller and Köhnken presented a systematic compilation of content characteristics that previous publications had named as being suitable for distinguishing between truthful and fabricated statements (Table 1). The fundamental assumption of this *Criteria-Based Content Analysis* (CBCA) is that statements based on memories of

Table 1. CBCA criteria (adapted from Steller & Köhnken, 1989)

General characteristics
1. Logical consistency
2. Unstructured production
3. Quantity of details
Specific contents
4. Contextual embedding
5. Descriptions 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 not comprehended
11. Related external associations
12. Accounts of subjective mental state
13. Attribution of perpetrator's mental state
Motivation-related contents
14. Spontaneous corrections
15. Admitting lack of memory
16. Raising doubts about one's own testimony
17. Self-deprecation
18. Pardoning the perpetrator
Offense-specific elements
19. Details characteristic of the offense

self-experienced events differ in quality from fabricated accounts (the so-called Undeutsch Hypothesis; Undeutsch, 1967). Steller and Köhnken compiled the CBCA criteria in order to operationalize this content quality. However, the number and type of content characteristics to be found in a statement are determined not only by the veracity of the statement but also by, for example, the nature of the event and the cognitive abilities of the witness – and hence this information needs to be taken into account as well. It was therefore pointed out that CBCA needs to be applied within a more comprehensive overall diagnostic procedure called *Statement Validity Assessment* (SVA; Köhnken & Steller, 1988; Steller, 1989).

The 25 years since Steller and Köhnken (1989) have seen a marked increase in knowledge about the distinction between experience-based and non-experienced-based statements. In the following, it will be argued that this newer knowledge strongly supports the SVA approach toward distinguishing between truthful and false statements on the basis of content characteristics. Moreover, new findings permit a better explanation of the underlying processes and, as a result, a better description of the overall diagnostic procedure.

There is also one major extension: Over the last 25 years, it clearly has been confirmed that repeated suggestive interviews performed with an expectation bias as well as explicit searches for until then nonretrievable memories can lead to subjective beliefs and, at times, to vivid mental percepts – that is, to false memories that have objectively

not occurred (e.g., Loftus, 2005). As will be shown, a non-critical application of CBCA may well lead to false results in these cases.

The Basic Psychological Issue in Statement Validity Assessment

The basic issue addressed here is whether or not a specific statement on an alleged offense is based on actual experience. Because witnesses' testimony refers to personally experienced events in the past, the relevant psychological question is whether one can find differences in behavior or statements when persons deliver experience-based versus non-experience-based reports.

Before addressing this question, it is necessary to define the category "non-experience-based." Deviation from the truth is not a uniform phenomenon, but one that may relate to a range of different psychological processes. Basically, one can describe two categories of non-experience-based reports: *intentional fabrications* and *accounts based on false memories*.

Both these categories break down into further subgroups. For example, an event may be fabricated in its entirety, something fabricated may be added to an actual event, or an actual event may be reported but transferred to another person who is falsely accused. In addition, false memories can form on the basis of both externally suggestive and autosuggestive processes (Steller, Volbert, & Wellershaus, 1993).

Although accounts are objectively not accurate in both of the above categories, there is one major difference: The witnesses differ in terms of their *psychological status*. Whereas lying witnesses are aware that they are deceiving, those making a statement based on a false memory are subjectively convinced of its truth. In this sense, their psychological status corresponds to that of a person giving true testimony. The psychological status of a liar, in contrast, differs from that of a truth teller. Hence, the differences to be found between experience-based and fabricated statements are probably not the same as the differences between experience-based statements and those based on false memories. Although it is acknowledged that witnesses may be located between these two poles (e.g., Von Hippel & Trivers, 2011), for the sake of clarity, the following will take a closer look at only these two basic constellations.

Experience-Based Versus Fabricated Accounts

There are two major differences between a truth telling and a lying witness: (a) The truth teller exhibits coherence between statement and belief, whereas a liar experiences a discrepancy between the two. (b) The statement of a truth teller is based on memory, whereas this is not (fully) the case for a liar.

There is a variety of theoretical proposals regarding what could form a basis for distinguishing between truthful and deceptive accounts (Vrij, 2008, for more detail). These proposals start with the aspect of deception and not that of missing memories, because they do not focus exclusively on memory-related forms of deception. For example, Zuckerman, DePaulo, and Rosenthal (1981) have pointed out that deception may lead to arousal (resulting in increased autonomous physiological reactions such as eye blinking or a higher pitched voice), the evocation of specific emotions (e.g., guilt that may lead to gaze aversion; or fear that may lead to increased physiological arousal with the aforementioned consequences in behavior), more marked control (resulting in, e.g., less spontaneous, more rigid behavior), or cognitive load (leading to, e.g., speech hesitations or longer response latencies).

A great number of studies have examined whether a person's behavior reveals cues that point to deception. However, meta-analyses (e.g., DePaulo et al., 2003) have shown hardly any correlations between behavioral cues and deception. In part, this may relate to the fact that situational conditions (e.g., the complexity of the lie, the motivation of the liar to get away with the lie, the stakes involved, or the interview style) vary greatly between studies (DePaulo & Morris, 2004; Vrij, 2008). Nonetheless, in general, the available findings indicate that behavioral cues for deception are, at best, only faint (Hartwig & Bond, 2011; Vrij, 2014; Vrij & Granhag, 2012).

Anxiety-Based Versus Cognitive Load Approach

Some authors have argued that empirical support for anxiety-based approaches is especially weak, making it seem more promising to concentrate on phenomena due to differences in cognitive load (Vrij & Granhag, 2012; see Frank & Svetieva, 2012, for a critical account).

Regardless of how far this conclusion may apply in general, anxiety-based approaches should be, in any case, of little help when trying to distinguish between statements based on experiences as a victim and fabricated testimony. In real forensic situations, both fabricated and experience-based testimonies may elicit arousal. It can also be assumed that experience-based reports on sexual abuse or other offences will also evoke negative emotions such as fear or shame.

In contrast, cognitive load is very significant for the issue addressed here. A real forensic situation confronts lying witnesses with a twofold task: They have to simulate an episodic memory while simultaneously concealing their deception. To achieve this, they must cope with a host of different tasks at the same time. They have to fabricate a complex statement without being able to draw on an episodic memory; and because they need to produce this statement within an interrogative setting, they have to be able to elaborate their fabricated account without contradictions and delay when faced with further questioning. At the same time, they continuously have to ensure that their deception will not be discovered. This means that they must attend to

the interviewer to find out whether they are still being believed (Buller & Burgoon, 1996). Moreover, liars must monitor their fabrications to match anything the interviewer knows or might find out. In addition, witnesses often have to testify not just once but repeatedly; and the intervals between interviews may be long – sometimes even lasting years. Because interviews are protocolled and these records serve as a basis for later interviews, lying witnesses have to take precise note of not only their original fabrications but also their spontaneous answers to further questions in order to reproduce a consistent statement at a later date. Witnesses giving experience-based testimony, in contrast, do not have to simulate any memories and they have no deception to conceal. They simply report from memory. Although this may well be stressful, it will be cognitively far less demanding.

Cognitive Scripts Versus Episodic Representations in Memory

Hence, one major difference between truthful and fabricated testimony is that truthful statements are based on memory whereas fabricated ones are not. Liars have to construct their statements from cognitive scripts (e.g., Schank & Abelson, 1977) containing general properties typical for that category of events. In contrast, the event-specific autobiographical representations that a truth teller draws on possess an episodic character and contain information on specific, spatiotemporally localizable events (e.g., Conway & Pleydell-Pearce, 2000). Depending on the actual experience in the particular case, there may well be visual, auditory, olfactory, spatial, and verbal information stored in memory; and some of this information might be unusual or even contrary to that in scripts (Köhnken, 1990).

The reality monitoring approach also points to such memory-related differences, although it does not refer to the aspect of deception. In their seminal paper, Johnson and Raye (1981) proposed that memories of externally derived experiences will contain more contextual, sensory, and semantic details than internally derived ones, whereas the latter will contain more references to cognitive processes at the time of encoding.

Experience-based statements should therefore contain more sensory and contextual information, be more strongly individual in character, and contain more details that would be inconsistent with or unrelated to cognitive scripts. This should result in what are generally more detailed, elaborate statements.

Strategic Self-Presentation

Liars pursue the goal of making a credible impression on their listener to render their false statements more convincing. In their *self-presentational perspective*, DePaulo and colleagues (DePaulo, 1992; DePaulo et al., 2003) argue that liars are typically less likely than truth tellers to take their credibility for granted. Although truth tellers are also keen

to appear credible, they typically do not expect this to require any special effort. Subsequently, liars are more concerned than truth tellers about the impression they make on others.

Attempts at strategic self-presentation consume mental resources, compromise performance, and lead to less elaborate accounts. Liars can also be assumed to avoid admissions of imperfection in their accounts or elements they believe will indicate deception (Köhnken, 1996, 2004).

Effects on Testimony

The aforementioned considerations show that relevant characteristics for determining the credibility of a statement are ones that (a) point to an episodic autobiographical representation in memory rather than a script-like construction of testimony and (b) show that witnesses are not making a particular effort to present themselves as credible communicators. Hence, these are not characteristics indicating deception, but ones pointing to *a relation to experience*. Accordingly, true statements – compared to fabricated ones – (a) are generally more elaborate and detailed; (b) contain more event-specific, sensory, contextual, individually relevant, script-deviant, or script-irrelevant details; (c) contain more details that would be believed to indicate deception from a lay perspective; and (d) contain more details that can be verified (see Volbert & Steller, 2009, for more detail).

Criteria-Based Content Analysis (CBCA)

CBCA criteria are a compilation of characteristics for appraising the quality of statements in the aforementioned sense. Although the ideas presented here were not formulated so explicitly in early publications on CBCA and SVA, these publications nonetheless always emphasized characteristics that (a) indicate a relation to experience (as opposed to emphasizing indicators of deception), (b) focus on content characteristics (and not behavioral ones), and (c) start off by viewing the statement as a result of cognitive performance (Köhnken, 1990, 1996).

A closer inspection of the CBCA criteria (e.g., Köhnken, 2004, for a detailed description) reveals that they contain characteristics referring mostly to either event-specific representations in memory versus script (e.g., *contextual embedding, unexpected complications during the incident, unusual details*), or self-presentation (e.g., *admitting lack of memory, raising doubts about one's own testimony, spontaneous corrections*). It also shows that the approach is based on the assumption that the testimony of a truth teller is more elaborated in general (e.g., *quantity of details*).

The *validity of the CBCA criteria* has meanwhile been examined in a host of laboratory studies as well as a few field studies. In summary, these studies confirm the hypothesized difference in the content quality of truthful versus fabricated statements. A recent meta-analysis (Werner,

Naefgen, Koppehele-Gossel, Banse, & Schmidt, 2014) showed that a random-effects model produced a large estimated effect size of $d = 1.005$ ($k = 45$, $n = 2,473$, 95% CI [.723, 1.287]) when the dependent variable was the total CBCA score.

Whereas this meta-analysis did not examine the significance of single CBCA criteria, the meta-analysis carried out by DePaulo et al. (2003) reported small to moderate effect sizes for *logical structure* ($d = -.25$) *quantity of details* ($d = -.30$), *admitting lack of memory* ($d = -.42$), and *spontaneous corrections* ($d = -.29$). For other CBCA characteristics (*contextual embedding, unusual details*), effect sizes took the expected direction but failed to attain statistical significance. However, it should be noted that these meta-analytical findings on single CBCA criteria were based on a maximum of only six studies. In addition, not all criteria were assessed, particularly those that occurred rather infrequently.

Vrij (2008) reviewed the proportion of overall correct classifications performed on the basis of CBCA characteristics in 24 studies. He found that they varied from 54% to 90% with an average total accuracy rate of 70.5%. The average accuracy rate for the truth (hit rate) was 70.8% (44–91%); and for lies (correct rejections), it was 71.1% (60–100%). However, the majority of these percentages were based on discriminant analyses that tend to produce overoptimistic estimates of classification rates. Nonetheless, an analysis based only on those studies taking the judgments of raters into account also revealed an overall accuracy rate of 71.8% (63–83%; five studies). CBCA raters seem to be able to identify truthful statements better than fabricated ones (truthful: 78–91%; fabricated: 60–74%), but this information was only given in three studies. Altogether these accuracy rates are markedly higher than the average accuracy rate of 54% (61% for truthful and 47% for deceptive reports) reported in Bond and DePaulo's (2006) meta-analysis of assessments that did not use any specific assessment procedure. Direct comparisons between naive raters and raters familiar with CBCA produced mixed results, although it has to be stressed that the trained raters in these studies had often received only a very short training (Vrij, 2008).

It should be noted that the findings in the laboratory research constituting the majority of validity studies were generally collected under conditions that deviate in a great number of ways from real forensic situations and place much lower cognitive load on their participants compared to witnesses giving testimony (often free narratives without questioning; frequently even written reports with no need to hide deception; no consequences when deception was discovered; no possibility of checking the information given; nearly always one-shot surveys). Given such a reduction in task demands, differences between truthful and fabricated statements due to variations in cognitive load or self-presentation strategy should be smaller than in real forensic situations.

Nonetheless, differences in CBCA characteristics between truthful and fabricated statements are still not large enough to warrant the use of CBCA criteria as a checklist-like deception (or better truth) detection tool

Table 2. Modified system of content characteristics (based on Niehaus (2008a))

Level	Autobiographic episodic memory versus script information	Strategic self-presentation
<i>Single characteristics</i>	Characteristics of episodic autobiographical memory, e.g.: <ul style="list-style-type: none"> – Contextual embedding – Spatial information – Temporal information – Description of interactions – Reproduction of conversations – Peripheral details – Sensory impressions – Emotions and feelings – Own thoughts – Personal implications – Perpetrator's mental state Script-deviant/script-irrelevant details, e.g.: <ul style="list-style-type: none"> – Unexpected complications – Unusual details – Related external associations Details not comprehended <ul style="list-style-type: none"> – Accurately reported details not comprehended 	Indications of memory-related shortcomings, e.g.: <ul style="list-style-type: none"> – Spontaneous corrections – Admitting lack of memory – Efforts to remember – Reality controls Questioning credibility, e.g.: <ul style="list-style-type: none"> – Raising doubts about one's own testimony – Raising doubts about one's own person Other problematic contents, e.g.: <ul style="list-style-type: none"> – Self-deprecation – Pardoning the perpetrator
<i>Statement as a whole</i>	<ul style="list-style-type: none"> – Reconstructability of the event – Vividness of event – Quantity of details – Unstructured production – Spontaneous supplementing 	

(Köhnken, 2004). Truthful and fabricated statements cannot be distinguished merely on the basis of the presence of CBCA criteria. As already mentioned in the introduction, content quality is influenced not only by the veracity of the statement but also by other (personal and contextual) variables such as the complexity of the event and the cognitive abilities of the witness. This information also has to be taken into account, as will be pointed out in more detail below.

Further Development of the System of Content Characteristics

It should be stressed that the set of CBCA criteria compiled by Steller and Köhnken (1989) made no claims to be exhaustive. The characteristics were meant as examples of content quality that could be utilized to answer the two underlying diagnostic questions: (a) Could a witness produce a testimony with this specific quality of content if it were not based on a real experience? (b) Would the witness produce a testimony with this specific quality of content if it were not based on a real experience?

As pointed out above, reality monitoring criteria assess memory-related characteristics of the differences between externally or internally derived accounts, although they do not consider the aspect of deception and strategic self-presentation (e.g., Masip, Sporer, Garrido, & Herrero, 2005). Sporer (2004) and Vrij (2008) have both already

pointed to major overlaps between reality monitoring and CBCA criteria. Indeed, some memory-related aspects are formulated in a more differentiated way in the reality monitoring criteria than in the CBCA criteria. For example, the latter do not assess sensory information very explicitly.

Niehaus (2008a) has presented a modified set of content-related characteristics that places greater emphasis on the underlying issues. She distinguishes characteristics pointing to the differences in the cognitive processes of liars and truth tellers from characteristics referring to the aspects of strategic self-presentation. The compilation by Niehaus can be extended further (Table 2). One first basic differentiation is that between (a) characteristics referring to the distinction between episodic memory and script information versus (b) characteristics referring to strategic self-presentation.

Memory-related characteristics can be further distinguished into three subcategories. The first is (a) details characterizing episodic autobiographical memories, that is, specific spatiotemporal and self-related information. It has to be considered that lying witnesses also have to refer to such details when trying to fake an episodic memory. However, it is assumed that liars will not succeed in presenting and integrating these details in such an elaborated way as truth tellers who draw on an actual memory. The second subcategory is (b) script-deviant details (e.g., *unusual details, unexpected complications during the incident*). These should occur rarely in the account of a lying witness,

because liars will probably not come up with the idea of integrating such information. Finally, (c) one characteristic (*accurately reported details not comprehended*; a characteristic found quite often in the statements about sexual abuse made by children who know little about sexuality) should not occur at all in a fabricated testimony, because it is impossible to fabricate beyond one's own understanding.

Three subcategories can also be distinguished among the characteristics referring to *strategic self-presentation*. The first is (a) characteristics indicating memory-related deficits (e.g., *spontaneous corrections*, *admitting lack of memory*). It is assumed that lying persons, who are only faking memory processes, will tend not to point to memory deficits in their testimony, because this would make giving testimony even more difficult; and moreover, indications of memory deficits could trigger skepticism in the interviewer. Truth tellers, in contrast, are actually remembering, and this process may well lead them to comment on their memory processes. The second is (b) characteristics assessing contents that cast doubt on credibility (*raising doubts about one's own testimony*, *raising doubts about one's own person*) and the third is (c) characteristics assessing other problematic contents (*self-deprecation*, *pardoning the perpetrator*).

These distinctions suggest that characteristics should be weighted differently. Arntzen pointed out already in 1970 that simple manifestations of characteristics are irrelevant, and that only quantitatively or qualitatively marked characteristics are important for credibility assessment. This should apply particularly to characteristics indicating episodic memory, which have to be expected in fabricated accounts as well – although in a less elaborated form. Script-deviant details, in contrast, should be very rare in fabricated testimony, and accurately reported details that have been not comprehended should not occur at all.

A study by Hommers (1997) underlines the assumption that a particular weight should be assigned to the latter characteristics. He showed that it was the characteristics *accurately reported details not comprehended*, *unexpected complications*, and *related external associations* that continued to distinguish between true and fabricated testimony among a group of good liars (defined as individuals whose fabricated statements contained a higher total CBCA score than their truthful accounts).

The approach proposed here is also supported by studies dealing with deception strategies. Niehaus and colleagues (Niehaus, 2008b; Niehaus, Krause, & Schmidke, 2005) presented participants with single content characteristics and asked them to report which ones they would integrate into a fabricated account and which ones they would avoid. Participants reported integrating *contextual embedding*, *descriptions of interactions*, *reproduction of conversation*, and *accounts of subjective mental state* into a fabricated testimony, but wanting to avoid *unusual details* and *superfluous details* as well as *spontaneous corrections*, *efforts to remember*, and *memory uncertainties*.

Furthermore, studies in which interviewees were informed about CBCA criteria showed that the statements of coached participants were assigned higher CBCA scores than those of uninformed participants. However, when

participants were given the general instruction to integrate CBCA characteristics but were free to choose which characteristics they used, significantly higher scores were found only on characteristics pointing generally to episodic memory (*contextual embedding*, *accounts of subjective mental state*, *attribution of perpetrator's mental state*, *reproduction of conversation*; Rutta, 2001). Other characteristics (*unusual details*, *unstructured production*, *accurately reported details not comprehended*, *unexpected complications*) were not incorporated significantly more frequently into the fabricated statement after training, even when the participants were explicitly instructed to integrate them (Wrege, 2004). In line with this, Vrij, Kneller, and Mann (2000) also found that although some CBCA characteristics could be augmented through training (*quantity of details*, *attribution of perpetrator's mental state*, *raising doubts about one's own testimony*), there were, however, no significant training effects for *unstructured production*, *unusual details*, and *superfluous details*.

Up to now, the least theoretical backing can be found for characteristics referring to other problematic contents (*self-deprecation*, *pardoning the perpetrator*). They are based on considerations of plausibility ("A lying witness who wrongly accuses someone will not exonerate him at the same time"). However, some empirical findings cast doubt on these characteristics. For example, an analysis of single CBCA characteristics showed that *self-deprecation* was the only characteristic to receive no empirical support at all (Vrij, 2008). It did not reveal higher scores for true compared to fabricated testimony in any study; and, in two out of ten studies, the criterion even appeared significantly more often in the fabricated statements. In addition, Niehaus (2003) found that although participants reported wanting to avoid objections to their own person or their own testimony in fabricated statements, they evaluated self-incriminations and exoneration of the accused positively. Nonetheless, when it was explicitly a false testimony on a sexual offence, participants reported that they would avoid integrating these contents into their testimony (Niehaus et al., 2005). These findings show that deception strategies depend on their context.

Statement Validity Analysis: A Diagnostic Approach for Single-Case Assessment

The central diagnostic question when performing a credibility assessment in a single case is whether *this particular witness* could have or would have produced this testimony on *this particular alleged event* if it were not based on real experience. This basic approach can already be found in the original publications on SVA. However, the Validity Checklist published at that time (Steller, 1989) may not have conveyed this basic approach sufficiently.

To answer the central question, it is necessary to estimate the possible impact of personal and situational variables on content quality. Some relevant variables will be briefly addressed in the following (see Volbert, 2010; Volbert, Steller, & Galow, 2010, for more detail).

Personal Variables

Age of the Interviewee

Statements by younger children are likely to contain fewer details than statements by older children and adults, because cognitive ability, command of language, and memory retrieval strategies mature throughout childhood, making it gradually easier to produce detailed accounts. Research has demonstrated that total CBCA scores are age dependent: With increasing age, children obtain higher total CBCA scores in both truthful and fabricated accounts (see Vrij, 2008, for more detail).

General Tendency to Narrate Autobiographical Experiences

This addresses interindividual differences in the ability to access specific autobiographical memory and to express oneself verbally. However, what is most important is the individual-specific tendency to communicate autobiographical experiences, for example, in terms of the amount of detail (Nahari & Vrij, in press).

Relevant Prior Experiences/Familiarity With the Event

If the interviewee is familiar with the event or if there have been prior experiences that are similar in content to that addressed in the statement, this facilitates the task of constructing a fabricated statement (e.g., Niehaus, 2001; Pezdek et al., 2004).

Personality Characteristics

Schelleman-Offermans and Merckelbach (2010) have shown that statements by high fantasy-prone individuals contain more CBCA characteristics than those by low fantasy-prone individuals. Vrij, Akehurst, Soukara, and Bull (2002) have found that in some age groups, CBCA scores correlated positively with social adroitness and negatively with social anxiety in both truthful and fabricated accounts. They also found a positive correlation between self-monitoring and CBCA scores in liars but not in truth tellers. General tendencies to dramatize are also relevant here (cf. Steller & Böhm, 2008).

Abilities to Deceive/Deception Strategies

It has to be assumed that some people are better liars than others (e.g., Vrij, Granhag, & Mann, 2010). How successfully a complex testimony can be constructed may depend, among others, on the intellectual capacity and the creative potential of the witness. It can also be assumed that persons who experience less emotion while lying need to concentrate less on controlling their emotions. This frees more cognitive resources for producing their statement compared to persons in whom false testimony triggers a greater

degree of anxiety or shame (cf. Vrij, Granhag, & Porter, 2010). Deception strategies are generally learned within social interaction and refined with the help of feedback from the social environment. Insofar, these also need to be taken into account in relation to the age and intellectual capacity of the witness (Niehaus, 2008b).

Willingness to Testify

Testimony should not be viewed as being identical to recall from memory. Reporting from memory is an active process in which decisions are made on how much information to disclose. If a witness is very unwilling to testify, this may result in poor content quality even when the testimony is based on actual experience.

Situational Variables

Complexity of the Event in Question

If the event in question is a simple occurrence that lasted only a few seconds, any statement on it will most likely not have a very high content quality – even if it does relate to a real experience. One can only describe CBCA criteria such as *unexpected complications* or *unusual details* when such elements were also present in the original event. However, this is generally not the case in simple events.

Time Interval Between Event in Question and Interview

If there is a long interval between the event and the interview, details may not be remembered well, and this can impact negatively on the content quality of testimony.

One-Off Versus Multiple Similar Events

Particularly sexual offences against children are frequently no single, one-off events. With multiple similar events, there is a tendency to form generic memory representations (e.g., Howe, 2000). As a consequence, specific details may no longer be remembered over the course of time (cf. Brubacher, Malloy, Lamb, & Roberts, in press).

Interview Technique

If predominantly specific questions are posed, no free narrative is possible, and content quality characteristics such as *unstructured production* or *superfluous details* cannot be produced. In contrast, open questions and invitations to free narratives lead to longer answers and a better content quality than specific questions. Craig, Scheibe, Raskin, Kircher, and Dodd (1999) as well as Hershkowitz (1999) have already shown that the differences between truthful and fabricated accounts increase in terms of CBCA scores when such an open interview technique is used. In a field study of

testimony from sexual abuse cases, Hershkowitz, Fisher, Lamb, and Horowitz (2007) found that the use of the NICHD protocol (Orbach, Hershkowitz, Lamb, Esplin, & Horowitz, 2000), a technique designed to encourage narrative accounts, led to a marked improvement in total accuracy rates (60% compared to 30% when the NICHD protocol was not used). Statements with a high probability of being true were able to be identified particularly well when the NICHD protocol was used (95% with compared to 38% without NICHD protocol). Statements with a high probability of not being based on experience were, nonetheless, hard to identify under both conditions, although, here as well, an advantage was found for the NICHD protocol conditions (24% compared to 12%).

Integrating the Findings

This information on personal and situational variables provides the basis for performing a quality-competence comparison (Steller, 2008) with which to conclude whether the witness was able to fabricate the statement in question or not. The assessment process concentrates on whether the content quality of the statement – taking the individual and contextual conditions into account – is good enough to conclude that the statement could not have been fabricated by this witness. If the content quality lies within a range that could well be fabricated by this particular witness, the fabrication hypothesis cannot be rejected for logical reasons. Nonetheless, this does not necessarily imply that the statement is fabricated. As pointed out above, there may be a variety of reasons for low content quality. As a result, these statements can be differentiated further with psychological methods (Volbert, 2010; Volbert & Steller, 2009). Hence the outcome of the quality-competence comparison is the identification of statements that cannot be reconciled with the fabrication hypothesis because of their high content quality. In contrast, statements with low content quality cannot be classified as fabricated, at least not when this low content quality can be explained by one of the above-mentioned reasons.

Some critics have complained that the process of integrating information fails to follow any formalized rules (Vrij, 2008). However, it is an intrinsic feature of an idiographic approach that diagnostic rules of inference are not accessible to standardization (Greuel, 2009), but this is the only way to take the complex individual combination of variables into account. Nonetheless, inferences must be justified explicitly in each individual case in order to appraise the persuasiveness of the argumentation; and inferences must be based on theoretically and empirically sound assumptions. This aspect will be taken up again below.

Experience-Based Accounts Versus Accounts Based on False Memories

If the fabrication hypothesis is falsified because of high content quality, this does not necessarily confirm the veracity

of the testimony. As pointed out above, it is also necessary to test whether it could be based on a false memory.

A host of studies have shown that suggestive processes can result in the formation of complex false memories and to detailed descriptions of events that are believed to be true although they actually never happened (e.g., Bruck & Ceci, 2009; Erdmann, 2001; Loftus, 2005; Loftus & Bernstein, 2005).

The firm belief that a specific event (in this context, a sexual abuse) has happened forms the starting point for the induction of false memories in children by a third person. In these cases, the child is initially silent and does not make any unsolicited or spontaneous statements about sexual abuse. Allegations emerge only after an adult suspects that something has occurred and starts to question the child. Whereas the child initially denies that something had happened, repeated questioning may lead eventually to a “disclosure.” The belief that a child has been a victim of sexual abuse is based mostly on an interpretation of behavioral symptoms. However, when no specific sexual abuse symptoms are present (e.g., Kendall-Tackett, Williams, & Finkelhor, 1993), a diagnosis of sexual abuse based solely on behavioral symptoms is always premature. The overinterpretation of behavioral symptoms is often combined with the assumption that children will normally not disclose sexual abuse and will deny if asked. Accordingly, the belief is that specific interview techniques are necessary to facilitate children’s disclosures. However, the interview techniques applied often transmit the interviewer’s belief and are therefore suggestive in nature. Such interviews are then shaped to elicit statements that are consistent with the interviewer’s a priori assumptions. Questions that might deliver inconsistent information are not asked; answers are interpreted in line with the a priori belief; expected information is reinforced; inconsistent information is ignored or interpreted within the framework of the initial hypothesis; and interviews are continued until the child provides information consistent with sexual abuse. For children who have not gone through the suspected experience, this creates a structurally unclear interview situation resulting in susceptibility for suggested information. Eventually the child may well come to report and incorporate the interviewer’s belief (see, e.g., Bruck & Ceci, 2012, Bruck, Ceci, & Principe, 2006; Volbert, 1999, for more detail).

Among adolescents and adults, suggestive processes frequently originate in a poor mental state of the person concerned. There is often a need to find an explanation for a complaint that is, however, mostly difficult to ascertain. Explanations in which recognizably external circumstances or even guilty third persons can be identified can have an alleviating effect (Stoffels, 2004). Assumptions that traumatic experiences are typically repressed or dissociated and can therefore not be recalled explicitly support the idea that such experiences could be the cause of one’s suffering. In this context, individuals often search explicitly for non-retrievable memories of traumatic experiences. An intensive and persistent preoccupation with such possible experiences can elicit images that may be considered falsely to be genuine memories because of their vividness, familiarity, and ease of recall. Although there are many

cases in which psychotherapists initiate and support such processes, they may also be autosuggestive (see McNally, 2003; Porter, Peace, Douglas, & Doucette, 2012; Volbert, 2004).

Distinguishing Between True and Suggested Statements

Although a statement based on a false memory is objectively not true, the person making it, unlike a lying witness, is subjectively convinced of its truth. This results in coherence between statement and belief, and, as a consequence, a strong similarity with the psychological status of a truth-telling witness. Unlike liars, witnesses testifying on the basis of a false memory are not constructing a fabricated account intentionally; nor are they preoccupied with concealing the deception and engaging in associated strategic self-presentation. What is different, however, is that one statement is based on a genuine memory; the other, on an internally generated idea. Whereas the former should function in line with findings from memory psychology, such psychological regularities may not hold for the latter.

Differences in Content Quality

For the reasons mentioned above, CBCA characteristics associated with cognitive load and self-presentation would seem to be less appropriate for distinguishing between truthful reports and accounts based on false memories.

In contrast to the large amount of empirical research on qualitative differences between truthful and fabricated statements, very few studies have addressed the issue of qualitative differences between truthful and suggested statements. Although Crotteau (1994) found significant differences on some CBCA criteria (*Logical consistency*, *Contextual embedding*, *Reproduction of conversation*, and *Unusual details*), *Superfluous details*, and *Admitting lack of memory* were unexpectedly more marked in suggested statements. The total accuracy rates on the basis of the total number of CBCA characteristics were poorer than the judgments of raters who were not familiar with CBCA, and this was particularly the case when identifying suggested statements. Erdmann, Volbert, and Böhm (2004; Erdmann, 2001) also found that suggested and experience-based descriptions by 1st-graders did not differ in terms of the total number of CBCA characteristics. A significant difference in line with the hypothesis that true testimony will be of a higher quality than non-experience-based testimony was found only for *Quantity of details*. Blandón-Gitlin, Pezdek, Lindsay, and Hagen (2009) compared true and suggested reports by adults under two conditions: One group of participants was convinced that the fictitious event had happened but that they had developed no visual memory of it (partial memory); the other group had developed a full-scale false memory (full memory). Whereas significant differences were still found between true descriptions and the statements of participants in the “partial memory” condition, true statements no longer differed significantly from

those of participants who had developed full-scale false memories. The results of studies that did not differentiate between true and suggested statements on the basis of CBCA criteria but used other content-related and verbal characteristics also point in the same direction: Bruck, Ceci, and Hembrooke (2002) used number of details, number of spontaneous namings, cohesion of the statement (use of temporal markers, reproduction of conversations), and statement elaboration (use of emotion-related expressions, use of adjectives and adverbs) to differentiate between experience-based and suggested descriptions in preschool children. They found that the experience-based and suggested statements became increasingly similar over the course of repeated interviews, and induced accounts eventually even contained more descriptive information than experience-based ones.

In summary, evidence in support of qualitative differences between experience-based and suggested statements is either very weak or even nonexistent once a full-scale false memory has developed. As a result, the presence of CBCA characteristics cannot be considered as indicating that an account is based on experience when it has been preceded by intensive suggestive processes.

Differences in the Formation and Evolution of Statements

However, the fact that true statements are based on episodic memory whereas accounts based on false memories are generated internally can still reveal differences. Significant events are generally remembered well, and this also applies to memories of traumatic experiences (see McNally, 2003; Porter et al., 2012; Volbert, 2011, for overviews). In addition, memories refer to experiences gained in the past that have come to an end and are, therefore, limited by experience. False memories, in contrast, are always discontinuous; that is, the alleged experiences are not “remembered” after the event in question but at a later point in time, and generally as a reaction to suggestive interviews or to an active search for suspected but currently nonretrievable memories of traumatic experiences. In addition, false memories are not real memories and may therefore also reveal patterns that do not correspond to findings from memory psychology – for instance, when memories of events in the first months of life are reported in detail. Moreover, false memories do not refer to events in the closed past, and can therefore always carry on developing further. Hence, it is necessary to perform a precise *reconstruction of the history of the formation and evolution of the statement* in order to ascertain or exclude any potential suggestive influences.

To ascertain potential *suggestive influences on children*, it is necessary to know the following: (a) whether a suspicion or an expectation was present before the child’s statement (*interviewer bias*); (b) whether the child either already imparted information on the relevant event at the first opportunity to talk about it or initially denied it and disclosed only after repeated interviews and the use of suggestive interview techniques (*biased interviews*); and

(c) whether there was an outcome-open clarification of a suspicion, or whether this procedure served only a confirmatory purpose – that is, information was interpreted only in terms of the initial hypothesis (*confirmation bias*).

To clarify *externally suggestive or autosuggestive influences in adolescents or adults*, it is necessary to ascertain whether (a) a *discontinuous (recovered) memory* is being reported; (b) there was a specific expectation that previously nonretrievable, traumatic experiences must have occurred (*expectation bias*); and (c) there has been an active search for suspected traumatic experiences (*explicit search for until then nonretrievable memories*; see, also, Porter et al., 2012; Stoffels & Ernst, 2002)

If significant externally suggestive or autosuggestive influences can be found in the history of a statement, it can no longer be ruled out that these have led to a false memory, and that this may form the basis for the statement in question. In this case, for the reasons reported above, possessing a high content quality does not confirm a statement's truth. Hence, high content quality does not deliver a falsification of the suggestion hypothesis.

Vice versa, it may be possible to find elements in the evolution of the statement or the statement itself that are hard to reconcile with a real experience. If, alongside a suggestive history, the statement contains elements that clearly run counter to findings in developmental or memory psychology (e.g., detailed reports on events that are supposed to have occurred during the first 2 years of life), this indicates that suggestion effects are not just possible but actually present (e.g., Steller, 1998; Volbert, 2010; see Volbert & Steller, 2009, for an overview).

Summary and Overall Assessment

It has been shown that no behavioral indicators or content characteristics exist that correlate with the veracity status of an account in the sense of nomological laws. In contrast, one can find content-related characteristics that tend to occur more often in experience-based than in fabricated accounts. However, the relationship between these criteria and the veracity status of a statement is not strong enough to judge veracity merely from the presence or absence of a single criterion or a certain set of criteria. On the one hand, this is because some experience-based statements reveal a relatively low content quality for various reasons. On the other hand, statements that are objectively not true but based on a false memory can also achieve a high content quality as shown above.

CBCA criteria may therefore well suggest that a testimony is not an actively constructed false accusation because, for example, the listing of *Superfluous details* or *Unusual details* may exceed the cognitive capacities of a lying person under the condition of high cognitive load. However, persons making a statement on the basis of a false memory can also give detailed reports on the basis of their mental ideas, because their lack of any intention to deceive makes them unconcerned about having to report details consistently in any follow-up interviews. As a result,

one can also frequently find *Superfluous details* and *Unusual details* in statements based on false memories (see Volbert, 2010; Volbert et al., 2010, for detailed accounts).

If severe suggestive influences can be found in the history of the statement, there is no known way of ruling out the possibility that they have determined the statement either completely or to a major extent. In contrast, if one has sufficient information on the history of the statement to exclude suggestive processes, the suggestion hypothesis can be falsified. It is then necessary to test the fabrication hypothesis. This calls for the application of a content analysis while taking into account individual competencies, individual prior experiences, and situational variables (complexity and frequency of the event in question, temporal aspects, type of interview) to test whether the content quality is sufficiently high to falsify the fabrication hypothesis.

Therefore, the principal diagnostic approach in SVA is to compare and contrast different models offering alternative explanations for the given data (Fiedler & Schmidt, 1999). This makes it necessary to test whether the statement in question can be explained in any other way than as being based on real experience, or whether such alternative explanations are incompatible with the available data. If all theoretically conceivable not true hypotheses are incompatible with the available facts, they can be dismissed, and one may accept the opposite hypothesis: the assumption that the statement is true (Köhnken, 2004). Insofar, in contrast to Rassin's (2000) claim that CBCA shows a truth bias, this methodological principle explicitly counteracts the use of a purely affirmative approach.

In sum, the basic assumption of SVA is that a testimony derived from the memory of an actual experience differs in content quality from a statement based on fabrication. This is because liars cannot fall back on an event-specific representation in memory; and, through their need to appear credible, they are more preoccupied with strategic self-presentation than truth tellers. As a result, liars have to handle a higher cognitive load and are hence unable to achieve the same quality of testimony as truth tellers. For SVA, content quality is influenced not only by the veracity of the statement but also by other (personal and contextual) variables, and these have to be taken into account when analyzing the single case. Of special importance are suggestive influences that might have preceded the statement in question, because these can lead to false memories that can, in turn, produce high-quality testimony. Hence, SVA is not a quick-to-apply checklist-like tool. Instead, it delivers a *comprehensive ideographic approach* enabling psychological experts to perform what is, in many cases, a crucial assessment of the credibility of a specific testimony.

Accordingly, conducting a SVA calls for more information than just the statement. In Germany, psychological experts are granted access to all files containing the results of investigations by the police, the public prosecutor, and the court; they can themselves interview the witness before the trial, and during the trial they can further question not only the witness whose statement credibility is being assessed but also other witnesses or the accused.

With regard to practical applications, it should be noted that the distinction between truthful and fabricated

statements is mostly not at the forefront of credibility assessments by psychological experts. In the majority of cases in which psychological experts were appointed in the last 25 years, at least in Germany, there has been some concern that the incriminating testimony might be based on suggestive processes. This underlines that SVA is far more than just CBCA (Volbert, 2008).

Suggestions for Future Research

There is considerable empirical support for the basic assumption of a qualitative difference between experience-based and fabricated statements. The total accuracy rates obtained using CBCA characteristics are far higher than those obtained without using a specific assessment system; CBCA assesses true and fabricated statements equally well; and no other assessment system achieves higher accuracy rates. However, the approximately 70% accuracy rates obtained solely on the basis of CBCA ratings alone do not suffice for an individual assessment. The assumption is that the accuracy rate can be improved by taking into account case-specific (personal and situational) conditions.

However, to date there is a lack of empirical research on the complete SVA procedure. There is, for example, an urgent need to test the level of agreement between SVA experts. For obvious reasons, it would also be highly interesting to test the validity of expert judgments on the basis of all relevant case-specific information. However, it regularly proves difficult to determine the ground truth in field studies independently from credibility assessments. Nonetheless, the validity of SVA expert judgments should at least be tested on the basis of selected cases in which additional information is available.

In the past, there has always been a strong tendency to view differences between truthful and fabricated statements from a general perspective and to broadly ignore differential aspects. To improve the diagnostic strategy in specific cases, it is desirable to gather more information on how individual characteristics (such as verbal creativity, intelligence, social anxiety, or social adroitness – Vrij et al., 2002) relate to the quality of both true and fabricated statements. Such studies should not only assess correlations between individual characteristics and high-quality fabricated statements but also focus on conditions under which low quality is to be found in truthful statements (Volbert & Lau, 2013).

If laboratory studies are to generalize to real forensic situations, there is a need for the best possible match with real task demands (i.e., giving testimony in a face-to-face situation in which deception has to be concealed; giving longer statements; not just giving free recall, but being questioned; being questioned repeatedly or at least being told that one will be questioned again). It should also be noted that persons who decide to make false allegations that could have detrimental outcomes for others may well differ markedly from persons who comply with instructions to give false reports in laboratory experiments. Hardly any studies have addressed this issue.

Conclusions

SVA is a method that focuses on aspects of the content of a statement and not on the behavior of the witness. It is also a method that does not refer to indicators of deception, but to indicators of memory and of a lack of strategic self-presentation. In 1989, when Steller and Köhnken first introduced CBCA, this focus could not be taken for granted. However, over the last 25 years, it has been confirmed in principle by research on deception detection. The validity of the set of CBCA characteristics compiled by Steller and Köhnken has also been confirmed in principle by empirical studies. However, on the basis of present-day knowledge, the underlying mechanisms can be explained better, and this also means that the potentials and limitations of this method can be described more clearly than it was the case 25 years ago. The current knowledge about false memories has pointed particularly to the limits of this method. Vice versa, the false memories problem has contributed to a better understanding of the underlying mechanisms.

Whereas there have been many studies on the quality of the content of testimony over the last 25 years, hardly any research has addressed the entire diagnostic strategy in which the content analysis is embedded in order to reach a decision on the single case based on taking individual competencies and case-specific conditions into account. This is due to the fact that the complex interplay of case-specific variables is hard to simulate in a laboratory. Nonetheless, future research needs to test the level of agreement and – as far as that is possible – also the validity of expert judgments on the basis of such information. Future research should also pay more attention to differential aspects and examine how individual-specific variables are linked to the content quality of a statement to further improve Statement Validity Analysis as a diagnostic approach to single-case assessment.

References

- Arntzen, F. (1970). *Psychologie der Zeugenaussage* [The psychology of testimony]. Göttingen, Germany: Hogrefe.
- Blandón-Gitlin, I., Pezdek, K., Lindsay, D. S., & Hagen, L. (2009). Criteria-based content analysis of true and suggested accounts of events. *Applied Cognitive Psychology, 23*, 901–917. doi: 10.1002/acp.1504
- Bond, C. F. J., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234. doi: 10.1207/s15327957pspr1003_2
- Brubacher, S. P., Malloy, L. C., Lamb, M. E., & Roberts, K. P. (2013). How do interviewers and children discuss individual occurrences of alleged repeated abuse in forensic interviews? *Applied Cognitive Psychology, 4*, 443–450.
- Bruck, M., & Ceci, S. J. (2009). Reliability of child witnesses' reports. In J. L. Skeem & S. O. Lilienfeld (Eds.), *Psychological science in the courtroom: Consensus and controversy* (pp. 149–171). New York, NY: Guilford.
- Bruck, M., & Ceci, S. J. (2012). Forensic developmental psychology in the courtroom. In D. Faust (Ed.), *Coping with psychiatric and psychological testimony* (6th ed., pp. 723–736). New York, NY: Oxford University Press.
- Bruck, M., Ceci, S. J., & Hembroke, H. (2002). The nature of children's true and false narratives. *Developmental Review, 22*, 520–554. doi: 10.1016/S0273-2297(02)00006-0

- Bruck, M., Ceci, S. J., & Principe, G. F. (2006). The child and the law. In K. A. Renninger, I. E. Sigel, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology, Vol 4: Child psychology in practice* (6th ed., pp. 776–816). Hoboken, NJ: Wiley.
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*, 203–242.
- Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review, 107*, 261–288. doi: 10.1037/0033-295X.107.2.261
- Craig, R. A., Scheibe, R., Raskin, D. C., Kircher, J. C., & Dodd, D. H. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. *Applied Developmental Science, 3*, 77–85. doi: 10.1207/s1532480xads0302_2
- Crotteau, M. L. (1994). *Can criteria-based content analysis discriminate between accurate and false reports of preschoolers? A validation attempt*. Unpublished master's thesis, Cornell University, Ithaca, NY, USA.
- DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 111*, 203–243. doi: 10.1037/0033-2909.111.2.203
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74–118. doi: 10.1037/0033-2909.129.1.74
- DePaulo, B. M., & Morris, W. L. (2004). Discerning lies from truths: Behavioral cues to deception and the indirect pathway of intuition. In P.-A. Granhag & L. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 15–40). New York, NY: Cambridge University Press.
- Erdmann, K. (2001). *Induktion von Pseudoerinnerungen bei Kindern* [Induction of pseudomemories in children]. Regensburg, Germany: Roderer.
- Erdmann, K., Volbert, R., & Böhm, C. (2004). Children report suggested events even when interviewed in a non-suggestive manner: What are its implications for credibility assessment? *Applied Cognitive Psychology, 18*, 589–611. doi: 10.1002/acp.1012
- Fiedler, K., & Schmidt, J. (1999). Gutachten über Methodik für Psychologische Glaubwürdigkeitsgutachten. Wissenschaftliches Gutachten für den Bundesgerichtshof [Evaluation of the methodology and criteria of psychological credibility assessments. Scientific evaluation for the German Federal Court of Justice]. *Praxis der Rechtspsychologie, 9*, 5–45.
- Frank, M. G., & Svetieva, E. (2012). Lies worth catching involve both emotion and cognition. *Journal of Applied Research in Memory and Cognition, 1*, 131–133.
- Greuel, L. (2009). Was ist Glaubhaftigkeitsbegutachtung (nicht)? Zum Problem der Dogmatisierung in einem wissenschaftlichen Diskurs [What is credibility assessment (not)? On the problem of dogmatizing within a scientific discourse]. *Kindesmisshandlung und Vernachlässigung, 12*, 70–89.
- Hartwig, M., & Bond, C. F. J. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*, 643–659. doi: 10.1037/a0023589
- Hershkowitz, I. (1999). The dynamics of interviews involving plausible and implausible allegations of child sexual abuse. *Applied Developmental Science, 3*, 86–91. doi: 10.1207/s1532480xads0302_3
- Hershkowitz, I., Fisher, S., Lamb, M. E., & Horowitz, D. (2007). Improving credibility assessment in child sexual abuse allegations: The role of the NICHD investigative interview protocol. *Child Abuse & Neglect, 31*, 99–110. doi: 10.1016/j.chiabu.2006.09.005
- Hommers, W. (1997). Die aussagepsychologische Krieriologie unter kovarianzstatistischer und psychometrischer Perspektive [CBCA criteria from the perspective of covariance statistics and psychometrics]. In L. Greuel, T. Fabian, & M. Stadler (Eds.), *Psychologie der Zeugenaussage. Ergebnisse der rechtspsychologischen Forschung* (pp. 87–100). Weinheim, Germany: Psychologie Verlags Union.
- Howe, M. L. (2000). *The fate of early memories: Developmental science and the retention of childhood experiences*. Washington, DC: American Psychological Association.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85. doi: 10.1037/0033-295X.88.1.67
- Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin, 113*, 164–180. doi: 10.1037/0033-2909.113.1.164
- Köhnken, G. (1990). *Glaubwürdigkeit* [Credibility]. Munich, Germany: Psychologie Verlags Union.
- Köhnken, G. (1996). Social psychology and the law. In G. R. Semin & K. Fiedler (Eds.), *Applied social psychology* (pp. 257–281). Thousand Oaks, CA: Sage.
- Köhnken, G. (2004). Statement validity analysis and the “detection of the truth”. In P.-A. Granhag & L. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 41–63). New York, NY: Cambridge University Press.
- Köhnken, G., & Steller, M. (1988). The evaluation of the credibility of child witness statements in the German procedural system. *Issues in Criminological & Legal Psychology, 13*, 37–45.
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361–366. doi: 10.1101/lm.94705
- Loftus, E. F., & Bernstein, D. M. (2005). Rich false memories: The royal road to success. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 101–113). Washington, DC: American Psychological Association.
- Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law, 11*, 99–122. doi: 10.1080/10683160410001726356
- McNally, R. J. (2003). *Remembering trauma*. Cambridge, MA: Belknap Press/Harvard University Press.
- Naefgen. (2013). *How valid are content-based statement veracity assessment techniques?* Unpublished master's thesis, University of Bonn, Germany.
- Nahari, G., & Vrij, A. (in press). Are you as good as me at telling a story? Individual differences in interpersonal reality monitoring. *Psychology, Crime & Law*. doi: 10.1080/1068316X.2013.793771
- Niehaus, S. (2001). *Zur Anwendbarkeit inhaltlicher Glaubhaftigkeitsmerkmale bei Zeugenaussagen unterschiedlichen Wahrheitsgehalts* [Applicability of CBCA criteria in statements of varying truth content]. Frankfurt am Main, Germany: Peter Lang.
- Niehaus, S. (2003). Auf der Suche nach der Wahrheit – intuitive Täuschungsstrategien als Hilfsmittel [In search of the truth: Intuitive deception strategies as aids]. In C. Lorei (Ed.), *Polizei & Psychologie. Kongressband der Tagung „Polizei & Psychologie“ am 18. und 19. März 2003 in Frankfurt am Main* (pp. 535–559). Frankfurt a. M., Germany: Verlag für Polizeiwissenschaft.
- Niehaus, S. (2008a). Merkmalsorientierte Inhaltsanalyse [Criteria-based content analysis]. In R. Volbert & M. Steller (Eds.), *Handbuch der Rechtspsychologie* (pp. 311–321). Göttingen, Germany: Hogrefe.
- Niehaus, S. (2008b). Täuschungsstrategien von Kindern, Jugendlichen und Erwachsenen [Deception strategies employed by children, adolescents, and adults]. *Forensische Psychiatrie, Psychologie, Kriminologie, 2*, 46–56.

- Niehaus, S., Krause, A., & Schmidke, J. (2005). Täuschungsstrategien bei der Schilderung von Sexualstraftaten [Deception strategies in the description of sex crimes]. *Zeitschrift für Sozialpsychologie*, *36*, 175–187. doi: 10.1024/0044-3514.36.4.175
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect*, *24*, 733–752. doi: 10.1016/S0145-2134(00)00137-X
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., . . . Brodie, L. (2004). Detecting deception in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology*, *89*, 119–126. doi: 10.1037/0021-9010.89.1.119
- Porter, S., Peace, K. A., Douglas, R. L., & Doucette, N. L. (2012). Recovered memory evidence in the courtroom: Facts and fallacies. In D. Faust (Ed.), *Coping with psychiatric and psychological testimony* (6th ed., pp. 668–684). New York, NY: Oxford University Press.
- Rassin, E. (2000). Criteria based content analysis: The less scientific road to truth. *Expert Evidence*, *7*, 265–278. doi: 10.1023/A:1016627527082
- Rutta. (2001). *Der Effekt von Hintergrundwissen über aussagepsychologische Methodik auf die inhaltliche Qualität von intentionalen Falschaussagen* [The impact of knowledge about CBCA on the content quality of fabricated accounts]. Unpublished diploma thesis, Free University Berlin, Germany.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Oxford, UK: Erlbaum.
- Schelleman-Offermans, K., & Merckelbach, H. (2010). Fantasy proneness as a confounder of verbal lie detection tools. *Journal of Investigative Psychology and Offender Profiling*, *7*, 247–260. doi: 10.1002/jip.121
- Sporer, S. L. (2004). Reality monitoring and detection of deception. In P.-A. Granhag & L. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 64–101). New York, NY: Cambridge University Press.
- Steller, M. (1989). Recent developments in statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 135–154). New York, NY: Kluwer/Plenum.
- Steller, M. (1998). Aussagepsychologie vor Gericht: Methodik und Probleme von Glaubwürdigkeitsgutachten mit Hinweisen auf die Wormser Mißbrauchsprozesse [Criteria-based statement analysis in court: Methods and problems of credibility evaluations in child sex abuse cases]. *Recht & Psychiatrie*, *16*, 11–18.
- Steller, M. (2008). Glaubhaftigkeitsbegutachtung [Credibility assessment]. In R. Volbert & M. Steller (Eds.), *Handbuch der Rechtspsychologie* (pp. 300–310). Göttingen, Germany: Hogrefe.
- Steller, M., & Böhm, C. (2008). Glaubhaftigkeitsbegutachtung bei Persönlichkeitsstörungen [Credibility assessment in cases of personality disorders]. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, *56*, 101–109. doi: 10.1024/1661-4747.56.2.101
- Steller, M., & Köhnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217–245). New York, NY: Springer.
- Steller, M., Volbert, R., & Wellershaus, P. (1993). Zur Beurteilung von Zeugenaussagen: aussagepsychologische Konstrukte und methodische Strategien [Constructs and methodological strategies in evaluating the credibility of witnesses' testimony]. In L. Montada (Ed.), *Bericht über den 38. Kongreß der Deutschen Gesellschaft für Psychologie in Trier 1992* (pp. 367–376). Göttingen, Germany: Hogrefe.
- Stoffels, H. (2004). Pseudoerinnerungen oder Pseudologien? Von der Sehnsucht, Traumaopfer zu sein [Pseudomemory or pathological lying? On longing to be a trauma victim]. In W. Vollmoeller (Ed.), *Grenzwertige psychische Störungen. Diagnostik und Therapie in Schwellenbereichen* (pp. 33–45). Stuttgart, Germany: Thieme.
- Stoffels, H., & Ernst, C. (2002). Erinnerung und Pseudoerinnerung: Über die Sehnsucht, Traumaopfer zu sein [Memory and pseudomemory: On the yearning to be a trauma victim]. *Der Nervenarzt*, *73*, 445–451. doi: 10.1007/s00115-002-1327-y
- Undeutsch, U. (1967). Beurteilung der Glaubhaftigkeit von Zeugenaussagen [Assessing the credibility of witnesses' testimony]. In U. Undeutsch (Ed.), *Handbuch der Rechtspsychologie* (Vol. 11, pp. 26–181). Göttingen, Germany: Hogrefe.
- Volbert, R. (1999). Determinanten der Aussagesuggestibilität bei Kindern [Determinants of children's suggestibility]. *Experimentelle und Klinische Hypnose*, *15*, 55–78.
- Volbert, R. (2004). *Beurteilungen von Aussagen über Traumata. Erinnerungen und ihre psychologische Bewertung* [Evaluation of testimonies on trauma. Memories and their psychological assessment]. Bern, Switzerland: Huber.
- Volbert, R. (2008). Glaubhaftigkeitsbegutachtung – mehr als merkmalsorientierte Inhaltsanalyse [Statement validity assessment – more than criteria-based content analysis]. *Forensische Psychiatrie, Psychologie, Kriminologie*, *2*, 12–19.
- Volbert, R. (2010). Aussagepsychologische Begutachtung [Credibility assessment]. In R. Volbert & K.-P. Dahle (Eds.), *Forensisch-psychologische Diagnostik im Strafverfahren* (pp. 18–66). Göttingen, Germany: Hogrefe.
- Volbert, R. (2011). Aussagen über traumatische Erlebnisse. Spezielle Aussagen? Spezielle Begutachtung? [Reports on traumatic experiences. Special memories? Special assessments?]. *Forensische Psychiatrie, Psychologie, Kriminologie*, *5*, 18–31.
- Volbert, R., & Lau, S. (2013). Unspezifität des autobiographischen Gedächtnisses bei Depressiven – Konsequenzen für die Aussagequalität [Overgeneralization of autobiographical memory in depressed patients: Implications for testimony content quality]. *Praxis der Rechtspsychologie*, *23*, 54–71.
- Volbert, R., & Steller, M. (2009). Die Begutachtung der Glaubhaftigkeit [Credibility assessment]. In H. Foerster & K. Dreßing (Eds.), *Psychiatrische Begutachtung* (pp. 693–728). München, Germany: Elsevier.
- Volbert, R., Steller, M., & Galow, A. (2010). Das Glaubhaftigkeitsgutachten [The evaluation of credibility]. In H.-L. Kröber, D. Dölling, N. Leygraf, & H. Saß (Eds.), *Handbuch der forensischen Psychiatrie* (pp. 623–689). Heidelberg, Germany: Springer.
- Von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, *34*, 1–16. doi: 10.1017/S0140525X10001354
- Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, *11*, 3–41. doi: 10.1037/1076-8971.11.1.3
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). New York, NY: Wiley.
- Vrij, A. (2014). Interviewing to detect deception. *European Psychologist*, *19*, 184–194. doi: 10.1027/1016-9040/a000201
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2002). Will the truth come out? The effect of deception, age, status, coaching, and social skills on CBCA scores. *Law and Human Behavior*, *26*, 261–283. doi: 10.1023/A:1015313120905
- Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, *1*, 110–117.

- Vrij, A., Granhag, P. A., & Mann, S. (2010). Good liars. *Journal of Psychiatry & Law*, 38, 77–98.
- Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, 11, 89–121. doi: 10.1177/1529100610390861
- Vrij, A., Kneller, W., & Mann, S. (2000). The effect of informing liars about criteria-based content analysis on their ability to deceive CBCA-raters. *Legal and Criminological Psychology*, 5, 57–70. doi: 10.1348/135532500167976
- Werner, V. A., Naefgen, C., Koppehele-Gossel, J., Banse, R., & Schmidt, A. F. (2014). *The validity of content-based assessment techniques to distinguish true and false statements: A meta-analysis*. Unpublished manuscript, University of Bonn, Germany.
- Wrege, J. (2004). *Der Einfluss von Hintergrundinformationen auf spezielle Glaubwürdigkeitsmerkmale* [The impact of prior knowledge about CBCA on selected CBCA criteria]. Unpublished diploma thesis, Free University Berlin, Germany.
- Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 1–59). New York, NY: Academic Press.

Received July 19, 2013

Accepted February 18, 2014

Published online July 24, 2014

About the authors



Dr. Max Steller is a professor emeritus for forensic psychology in Berlin, Germany, who has been involved in refining, systematizing, and researching statement validity assessment.



Prof. Dr. Renate Volbert is a research fellow at the Institute of Forensic Psychiatry, Charité, Berlin, Germany and a Professor in the Department of Psychology, Free University Berlin, Germany. Her main research interests are: Credibility assessment, suggestibility, secondary victimization, and false confessions.

Renate Volbert

Institut für Forensische Psychiatrie
Charité – Universitätsmedizin Berlin
Oranienburger Str. 285
13437 Berlin
Germany
Tel. +49 30 8445-1419
Fax +49 30 8445-1440
E-mail renate.volbert@charite.de