

Detecting Deception in Children: An Experimental Study of the Effect of Event Familiarity on CBCA Ratings

Iris Blandon-Gitlin,^{1,3} Kathy Pezdek,¹ Martha Rogers,² and Laura Brodie²

The CBCA is the most commonly used deception detection technique worldwide. Pezdek et al. (2004) used a quasi-experimental design to assess children's accounts of a traumatic medical procedure; CBCA ratings were higher for descriptions of familiar than unfamiliar events. This study tested this effect using an experimental design and assessed the joint effect of familiarity and veracity on CBCA ratings. Children described a true or a fabricated event. Half described a familiar event; half described an unfamiliar event. Two CBCA-trained judges rated transcripts of the descriptions. CBCA scores were more strongly influenced by the familiarity than the actual veracity of the event, and CBCA scores were significantly correlated with age. CBCA results were compared with results from other measures. Together with the results of K. Pezdek et al. (2004) these findings suggest that in its current form, CBCA is of limited utility as a credibility assessment tool.

KEY WORDS: credibility; deception in children; child sexual abuse.

In child abuse cases, evaluating the veracity of the accusations is particularly difficult because in most situations no physical evidence exists, and the alleged perpetrator and the child victim are the only witnesses to the crime. The steady increase in allegations of child abuse (Sedlak & Broadhurst, 1996), combined with the difficulty evaluating children's reports (e.g. Ceci & Bruck, 1993), has increased the need for researchers to develop valid and reliable tools for assessing the veracity of children's accusations. The Statement Validity Assessment (SVA) technique is a comprehensive system developed in Germany to evaluate the veracity of written statements (see Raskin & Esplin, 1991a for a discussion of the SVA). The most highly researched aspect of the SVA system is its core component, the Criterion-Based Content Analysis (CBCA), which is reported to be the most widely used veracity

¹Department of Psychology, Claremont Graduate University, Claremont, California.

²Tustin, California.

³To whom correspondence should be addressed at Department of Psychology, Claremont Graduate University, 123 Eighth St., Claremont, California 91711; e-mail: iris.blandon-gitlin@cgu.edu.

Table 1. Content Criteria for Statement Analysis.
Reprinted From Raskin and Esplin (1991a, p. 279)

General characteristics
1. Logical structure
2. Unstructured production
3. Quantity of details
Specific contents
4. Contextual embedding
5. Interactions
6. Reproduction of speech
7. Unexpected complications
8. Unusual details
9. Superfluous details
10. Accurately reported details misunderstood
11. Related external associations
12. Subjective experience
13. Attribution of the accused's mental state
Motivation-related contents
14. Spontaneous corrections or additions
15. Admitting lack of memory or knowledge
16. Raising doubts about one's own testimony
17. Self-deprecation
18. Pardoning the accused

assessment technique worldwide (Vrij, Akehurst, Soukara, & Bull, 2002). In this study we used an experimental paradigm to test how effectively the CBCA discriminates between true and fabricated reports of familiar versus unfamiliar events.

The development of the CBCA was based on the assumption that statements derived from memory of a self-experienced event will differ qualitatively from statements that are based on fantasy or suggestion (Steller, 1989). From this notion 18 criteria were generated that are considered to be characteristics of truthful statements. Table 1 includes a list of the 18 CBCA criteria, organized into three major categories. The main category, General Characteristics, is based on the premise that a report of a self-experienced event will be logical and coherent (criterion #1), will have digressions or shifts in focus (criterion #2), and will contain a wealth of details (criterion #3). The other two categories, Specific Contents of the Statement and Motivation-related Contents, refer to cognitive and motivational factors that are likely to be present in true statements (Raskin & Esplin, 1991b). An account is considered likely to be true if a substantial number of the 18 criteria are present, with the first three criteria being necessary but not sufficient. Currently, however, no formalized system exists to weight different criteria or to obtain absolute cut-off scores that can be used to designate an account as true or false.

Both field and laboratory studies have been conducted to test the validity of the CBCA. In the typical field study, real cases of alleged sexual abuse are grouped into "confirmed" cases and "unconfirmed" cases based on the extent to which independent corroborating evidence is available (for example see Boychuck, 1991; Craig, Scheibe, Raskin, Kircher, & Dodd, 1999; Lamb et al., 1997). In the typical laboratory study, veracity is manipulated by asking children or adults to lie or tell the truth about target events (for example see Landry & Brigham, 1992; Tye, Amato, Honts, Devitt, & Peters, 1999; Yuille, 1988; Zaparniuk, Yuille, & Taylor, 1995). In both

field and laboratory studies, CBCA trained judges evaluate the written transcripts and make predictions as to which are true accounts and which are false. The results have provided mixed support for the utility of the CBCA in assessing children's truthfulness.

In addition, there are forensically relevant constraints on the utility of the CBCA. For example, Vrij, Akehurst, Soukara, and Bull (2002) examined the CBCA scores for descriptions of adults and children for an incident in which important information was erased from a blackboard. Participants lied or told the truth about the reported event and they were coached or not regarding the CBCA criteria. True statements received significantly higher CBCA scores than false statements, however, there were no differences between CBCA-informed liars and participants who reported the truth.

Another constraint on the utility of the CBCA, and one upon which the current study was based, was recently reported by Pezdek et al. (2004). In this study, two CBCA-trained judges evaluated transcripts of reports made by children who had previously undergone a stressful and invasive medical procedure called Voiding Cystourethrogram Fluoroscopy (VCUG). This procedure is painful, involves intrusive forced genital contact, and is not an elective procedure for the child. On the basis of a survey of practitioners, Pezdek et al. (2004) reported that the VCUG procedure approximates in many ways the circumstances involving sexual abuse, the situation for which CBCA was developed.

Three groups participated in the Pezdek et al. (2004) study. The relative unfamiliar group had received the VCUG only once. The relative familiar group had received the VCGU procedure more than once. The control group had received a genital exam and not the VCUG. CBCA scores were significantly higher for children's descriptions in the relative familiar condition ($M = 10.67$) than in the relative unfamiliar condition ($M = 7.68$). The control group received ratings similar to the relative familiar group ($M = 9.92$). Thus, CBCA scores are affected by the familiarity of the event being described; accounts of familiar events are more likely to be considered true than are accounts of unfamiliar events. This is a forensically important finding because it suggests that in real-world cases, to the extent that a child is familiar with and has available information about an event, their account of an incident is likely to have characteristics that produce a high CBCA score, indicating that the event did occur, whether this is true or not.

This research is an extension of the Pezdek et al. (2004) study. There are two limitations in the previous study. First, although the target event (the VCUG) was an appropriate one for generalizing to real-world cases of sexual abuse, children were not, of course, randomly assigned to conditions. Thus, it cannot be concluded that difference in CBCA scores among conditions were caused by the degree of familiarity with the VCUG. We argue that prior to generalizing the results of any research study to real-world applications, it is important to conduct experimental research to assess the role of the factors under consideration. Second, because in all three conditions in the Pezdek et al. (2004) study children described a true event, the study did not assess of the joint effects of familiarity and veracity on CBCA scores. Because the CBCA was developed to discriminate between truthful and deceptive statements, veracity should be part of any

assessment of the effectiveness of the CBCA. This study responds to these two concerns.

This study used a completely randomized factorial design in which familiarity was crossed with the truth-value of the event described. Children were asked to describe an event that did occur (true condition) or not (false condition). For half of the children the event was familiar, and for the other half the same event was unfamiliar. Two experienced CBCA-trained judges, blind as to the purpose of the study, rated verbatim transcripts of the children's descriptions. We hypothesized that true statements would receive higher scores than false statements, and familiar reports, regardless of truth-value, would receive higher CBCA scores than unfamiliar reports of the same event. No predictions were made about the interaction.

In a review of the false memory research Pezdek and Taylor (2000) reported that true accounts could be discriminated from false accounts on the basis of three memory qualities—clarity, confidence, and verbosity. The main purpose of this study is to assess the effects of familiarity and truth-value on CBCA scores. In addition, however, this study provides an opportunity to assess the role of these three other memory measures in discriminating between true and false events.

METHOD

Participants and Design

The 94 children who volunteered to participate in this study were from fourth, fifth, and sixth grades and attended a public school in a Metropolitan area of Los Angeles, California. The age range was 9–12 years ($M = 10.5$, $SD = 1.0$). The study involved a 2 (familiar vs. unfamiliar) \times 2 (true vs. false) between-subjects design. The distribution of participants' gender and the number of children in each of the four conditions was approximately equal. Although the exact ethnicity of the participants in this study was not obtained, in the participating school the majority of children were of ethnicities other than white. About 78% of the students in the school qualify for free and reduced priced meals, indicating that the majority of children were from lower income households. Eighty percent of the children spoke English at home. Whether children spoke English at home was not significantly correlated with CBCA scores, $r_{pb} = .19$. Indicating that English proficiency was not related to CBCA scores.

Materials and Procedure

The target event was the act of sewing a button on a shirt. Pilot tests showed that this target event is unfamiliar to most children younger than 12 years of age, and this was confirmed in exit interviews in this study. The procedure consisted of three phases. In the *familiarization phase*, children in the familiar condition received 1-hr training session on the proper procedure for sewing a button on a shirt. This involved the alignment of the button at the buttonhole, marking the place where the button was to be positioned, threading the needle, knotting the thread and starting the sewing from the inside of the shirt. During the familiarization phase, materials

were provided for children to practice the task. An experimenter (Experimenter A) who was not seen again by the participants, provided this training in the students' classroom (room A). Children in the unfamiliar condition were not in the classroom during the familiarization phase and did not receive the sewing training. Approximately half of the children were assigned to the familiar condition and half to the unfamiliar condition.

In the *performance phase* the following day, children in the true condition performed a series of tasks, including solving a puzzle, arranging cards in numerical order, and the target event of sewing a button on a shirt. In the false condition the children performed the same series of tasks as children in the true condition with the exclusion of the target event. These activities were directed by a different experimenter (Experimenter B) in a different location (room B) from the familiarization phase.

Two or three days later, in the *recall phase*, Experimenter B returned with each child to room B. Thus, the recall phase and the performance phase were conducted by the same experimenter and in the same room, and these were different from the experimenter and the room used in the familiarization phase. In the recall phase, the child was asked to describe the specific sequence of events that occurred with Experimenter B in room B during the performance phase. The children's statements were audiotaped. All children were asked to be as detailed as possible so that the person listening to their descriptions could understand what the children did. To provide an example of the level of detail that was expected in their reports, children were read a sample narrative about a different event, a trip to the library.

Children in the false condition were told that the sequence of events they performed was supposed to have included sewing a button on a shirt, however, the researcher accidentally forgot to have them do this. Experimenter B told them to provide a description of sewing a button on a shirt in their narrative of what occurred during the performance phase (in room B). Children received instructions to describe the event in as much details as possible and were also told not to reveal that the experimenter had forgotten to have them actually sew on the button.

When the children were asked to provide a description of the specific sequence of events that occurred with Experimenter B in room B two or three days prior, half were describing a true event and half were describing a false event. It is important to note that the children were not being asked to provide a general description of the task of sewing on a button. Rather, they were asked to describe the specific details of the task that they had performed in room B with Experimenter B. Thus, in this study the CBCA was used to discriminate between children's accounts of a specific event that either did or did not occur, and the effect of prior knowledge and experience on the children's descriptions of this specific episode was assessed.

Each interview started with a general open-ended question, "Tell me all of the things that you did the day before yesterday when we met." When each child stopped talking, several times they were asked, "Do you remember anything else that you did?" Then the interviewer prompted each child with more focused questions (e.g., "Tell me all the steps that you took to sew the button on the shirt"). The goal of the prompts was to obtain as much detail as possible from each child without being suggestive or leading. Children in all four conditions followed the same test

procedures and received the same prompts. Upon completion of the recall phase, children rated how confident they were of their memory for the sewing event on a 5-point scale (1 = *unsure*, 5 = *absolutely sure*), and the clarity of their memory for the target event on a 5-point scale (1 = *not very clear*, 5 = *very clear*).

CBCA Ratings of Data

The interviews were transcribed verbatim, and copies were provided to two CBCA-trained judges. Both raters are clinical forensic psychologists who were extensively trained by one of the main proponents of the technique, David Raskin. This training included a 3-day seminar that involved analyses and veracity assessment of multiple transcripts. One of the raters received additional training from Dr. Raskin and served as a rater in a number of research projects. She also collaborated with him in several legal cases involving application of the CBCA. The other rater used the CBCA in her dissertation research for which Dr. Raskin was a contributing member. Both raters also use the CBCA in their research and clinical practices.

The judges knew that the target event was sewing a button on a shirt, but they were unaware of the purpose of the study. To minimize any influence of children's descriptions of the true nontarget activities on CBCA scores, judges were instructed to focus their analyses only on the target event. Both judges first practiced with sample transcripts from pilot participants and discussed their disagreements. Subsequently, they independently rated each of the transcripts over a period of a few weeks. Judges were provided with a coding sheet that included a list of 16 CBCA criteria and a brief definition of each. Criteria #17 and #18 were not used since they are not applicable to the target event in this study.⁴ Following the procedure of Craig et al. (1999) and Lamb et al. (1997), each judge rated each of the 16 criteria either 1 (criterion present) or 0 (criterion absent). Each judge's score was the sum across all criteria present in each participant's transcript (range of possible scores: 0 to 16). The correlation between the two judges' ratings across all participants was $r = .90$, $p < .01$. Because this correlation was significant, a mean CBCA rating for each transcript was computed based on the two judges' ratings. The dependent variable was the mean CBCA score for each transcript. An item-by-item analysis using Cohen's Kappa revealed that inter-judge agreement on each of the 16 criteria ranged from .85 to 1.0 with a mean agreement of .95.

RESULTS

Two ratings were obtained by each judge for each transcript: (a) ratings of the unprompted recall data, and (b) total recall ratings (the sum of the ratings for both

⁴After reviewing a subset of our transcripts, Dr. Phillip Esplin suggested that all of the original 18 CBCA criteria be used with the exception of criteria #17 and #18. Although Kohnken, Schimossek, Aschermann, and Hofer (1995) tested an additional five criteria for inclusion in the CBCA, their own support for including these additional criteria is weak. From their research they concluded, "Does this mean that the additional content characteristics should be included in the list of CBCA criteria? Such a step would certainly be premature, and we would not recommend it" (p. 681).

unprompted and prompted recall data). Only the analyses of unprompted data will be reported here, as the pattern of results was similar for ratings of unprompted and total recall, $r = .86$, $p < .01$.

Differences Between the Groups on Total CBCA Scores

The main purpose of this study was to test the hypothesis that reports of familiar events, regardless of truth-value, would receive higher CBCA scores than would reports of unfamiliar events. Higher CBCA scores indicate that the event was more likely to have occurred. To test this hypothesis, a 2 (familiar vs. unfamiliar) by 2 (truth vs. false) between-subjects analysis of variance (ANOVA) was performed. Only familiarity was significantly related to mean CBCA scores, $F(1, 90) = 9.51$, $p < .01$, $\eta^2 = .10$. Children in the familiar condition, $M = 5.12$ ($SD = 1.76$), received higher CBCA scores than did children in the unfamiliar condition, $M = 4.01$ ($SD = 1.73$). Although transcripts in the true conditions received slightly higher CBCA scores, $M = 4.81$ ($SD = 1.79$), than those in the false condition, $M = 4.32$ ($SD = 1.85$), the main effect of truth-value was not statistically significant, $F(1, 90) = 1.88$, $p = .17$, $\eta^2 = .02$. The interaction did not approach significance, $F(1, 90) < 1.00$, $\eta^2 = .00$.

Differences Between the Groups on Each of the Three Categories of CBCA Criteria

A multivariate analysis of variance (MANOVA) was conducted to assess the effects of truth-value and familiarity on each of the three categories of CBCA criteria. These included General Characteristics (criteria 1–3), Specific Contents (criteria 4–13), and Motivation-Related Contents (criteria 14–16). For each participant, the judges' ratings for the criteria were summed within each of these three categories. These three sums served as the three dependent measures in the MANOVA.⁵

The overall MANOVA that included the three categories of CBCA criteria was only statistically significant for familiarity, $F(3, 88) = 4.22$, $p < .01$, $\eta^2 = .13$. This effect was largely due to the significant contribution of the General Characteristics category, $F(1, 90) = 10.66$, $p < .01$, $\eta^2 = .11$. The Specific Contents category was marginally significant, $F(1, 90) = 3.42$, $p < .07$, $\eta^2 = .04$. The Motivation-Related category was not statistically significant. There was no effect of truth-value, $F(3, 88) = 1.24$, $p > .05$, $\eta^2 = .04$ and no interaction, $F(3, 88) = 1.01$, $p > .05$, $\eta^2 = .03$. These results suggest that the familiar and unfamiliar groups could be differentiated on the basis of the criteria in both the General Characteristics category and to some extent, those in the Specific Content category of the CBCA. The true and false groups could not be differentiated on the basis of any of these three CBCA categories.

⁵The complete MANOVA analyses and tables are available from the first author upon request. We opted to sum scores within each category of CBCA to assess the relative contribution of the criteria instead of using the 16 individual criteria as dependent variables. This procedure was employed because for the large majority of participants many of the criteria received scores of 0; thus, there were no differences among groups to be accounted for. Lamb et al. (1997) and others have reported similar findings. This limitation was avoided in the MANOVA's reported in this paper by summing across the criteria within each of the three CBCA categories and using these sums in as dependent variables in the analyses.

CBCA Score and Age

As in other CBCA studies (e.g. Buck, Warren, Betman, & Brigham, 2002; Craig et al., 1999; Lamb et al., 1997), we assessed the relationship between age and CBCA scores. Despite the fact that the age range for participants in this study was relatively restricted (9–12 years), there was a significant correlation between age and CBCA scores, $r = .29$, $p < .01$; transcripts of older children received higher CBCA scores than did transcripts of younger children.

Analysis With Other Memory Measures

Additional 2×2 ANOVA's were conducted on each of the three false memory measures: clarity, confidence and verbosity. There was a main effect of both familiarity and truth-value on confidence. Children in the familiar condition were more confident of their memories for the event, $M = 3.48$ ($SD = 1.53$), than were children in the unfamiliar conditions, $M = 3.0$, ($SD = 1.62$); $F(1, 89) = 6.77$, $p = .01$, $\eta^2 = .07$. Similarly, children in the true conditions rated their confidence significantly higher, $M = 4.52$ ($SD = .65$) than children in the false conditions, $M = 1.89$ ($SD = 1.07$); $F(1, 89) = 221.78$, $p < .01$, $\eta^2 = .72$. The interaction did not approach significance, $F(1, 89) = 1.43$, $p = .235$, $\eta^2 = .02$.

For the measure of clarity, only the main effect of truth-value was significant. Children in the true condition rated their memory for the event as clearer ($M = 4.50$, $SD = .72$), than those in the false conditions ($M = 1.68$, $SD = .96$), $F(1, 88) = 252.92$, $p < .01$, $\eta^2 = .74$. The difference between the familiar ($M = 3.23$, $SD = 1.67$) and the unfamiliar ($M = 3.07$, $SD = 1.63$) conditions was not significant, $F(1, 88) < 1.00$, $\eta^2 = .00$, nor was the interaction, $F(1, 88) < 1.00$, $\eta^2 = .00$.

Verbosity scores were obtained by counting the number of different idea units recalled by each child. Verbosity scores were significantly higher for children in the familiar condition ($M = 8.77$, $SD = 2.69$) than the unfamiliar condition ($M = 5.80$, $SD = 2.59$), $F(1, 90) = 29.67$, $p < .01$, $\eta^2 = .25$. The mean verbosity score in the true condition ($M = 7.69$, $SD = 2.77$) and the false condition ($M = 6.91$, $SD = 3.25$) did not significantly differ, $F(1, 90) = 2.08$, $p = .15$, $\eta^2 = .02$, nor was the interaction significant, $F(1, 90) < 1.00$, $\eta^2 = .00$. The similarity of the findings with both CBCA scores and verbosity is not surprising given that verbosity is one of the general characteristics (criterion number 3) of the CBCA.

The matrix of intercorrelations among the four dependent measures and age of the participants is included in Table 2. CBCA scores are significantly correlated with verbosity scores, $r = .38$, $p < .01$, but not with confidence ($r = .12$) or and clarity ($r = .15$).

DISCUSSION

This study tested if CBCA scores are affected by the familiarity of the event being described using an experimental task in which participants were randomly assigned to conditions, and the veracity of the event was varied—the event described was true for one group and false for another. Pezdek et al. (2004) reported that CBCA scores are significantly affected by event familiarity. However their study

Table 2. Intercorrelations Among the Four Dependent Variables and the Age of Participants

Variable	CBCA	Confidence	Clarity	Verbosity	Age
CBCA					
<i>r</i>	—				
Confidence					
<i>r</i>	.12	—			
Clarity					
<i>r</i>	.15	.86**	—		
Verbosity					
<i>r</i>	.38**	.22*	.20*	—	
Age					
<i>r</i>	.29*	-.10	-.16	.22*	—

p* < .05. *p* < .01.

was a quasi-experimental design and the events described were all true. The results of this study corroborate and extend the findings of Pezdek et al. (2004). CBCA scores were significantly affected by the familiarity of the event recalled but not whether the event was true.

Children in the false condition in this experiment were asked to pretend that they had sewn a button on a shirt as part of the sequence in the performance phase and then to provide in their description, details of the fabricated event. The finding that the effect of familiarity overwhelmed the effect of truth-value on CBCA scores suggests that information acquired from prior knowledge and experience may play a more dominate role in children’s descriptions than information retained from the specific episode being recalled. Because the CBCA was designed to discriminate between children’s accounts of true events and fabricated events, and it is used for this purpose in courts of law, the shortcomings identified in this study are forensically relevant and suggest the need for revision.

Why do familiar events receive higher CBCA scores than unfamiliar events? The experience of a specific event is not stored as an isolated unit in memory. Rather, when a specific event is encoded into memory, related script-relevant knowledge is activated that aids in the comprehension of the event. This script-relevant knowledge includes generic information about the event acquired from prior knowledge and experience, and is stored in memory along with information retained from experiencing the event directly. By definition, individuals have more script-relevant knowledge in memory for familiar events than for unfamiliar events. Thus, children’s accounts of familiar events are likely to contain more information and have a more coherent structure than are their descriptions of unfamiliar events. As a consequence, CBCA scores would be higher for familiar events than for unfamiliar events. This is similar to the role of event plausibility hypothesized by Pezdek, Finger, and Hodge (1997) and Pezdek and Hodge (1999) to account for the increased probability of planting false events in memory.

In light of the easy access to sexually explicit material in our culture, and the commonness of conversations about sexual topics, many children have knowledge of sexual behavior that was acquired indirectly from these sources. A child’s description of an alleged incident of sexual abuse may include information that has

been obtained from other sources. The results of this study, along with those of Pezdek et al. (2004), suggest that to the extent a child is familiar with the alleged incident, their account is likely to have characteristics that produce a high CBCA score, indicating that the event did occur, whether it is true or not.

Several researchers have expressed concerns about the effects of age and developmental level on CBCA scores (Buck et al., 2002; Lamers-Winkelmann & Buffing, 1996; Lamb et al., 1997; Raskin & Esplin, 1991b). The results of this study provide additional support for this concern. Despite the restricted age range in this study (9–12 years old), there was a significant correlation between age and CBCA scores ($r = .29$).

This experiment also assessed an alternative set of veracity assessment measures. In a review of the false memory research, Pezdek and Taylor (2000) reported three qualities of memories reported to effectively discriminate between accounts of true and false events—the verbosity of the description of the event and self-ratings of the confidence and clarity of the memory. In this study the effects of familiarity and truth-value on these three measures were assessed. True and false accounts could be discriminated on the basis of both confidence and clarity ratings, suggesting that when children fabricate an account of an event, they are aware that the fabricated event did not occur. Even if children's descriptions of fabricated events have characteristics that produce high CBCA scores, children's memories for fabricated events are less clear and are less confidently held. Additionally, the correlations of age with confidence and clarity were not significant. Together these results suggest that these two memory qualities show promise as components of an effective veracity assessment measure.

On the other hand, verbosity scores were affected only by the familiarity and not the truth-value of the event. This finding was not surprising given that verbosity is comparable to the “quantity of details” criterion of the CBCA. This result provides further support for the conclusion that information acquired from prior experience may play a more dominant role in affecting children's accounts than information retained from the specific episode being recalled.

The results of this study, along with the findings of Pezdek et al. (2004), suggest that children's accounts of an event are affected to a greater extent by prior knowledge than by actual experience, and that CBCA scores assigned to children's accounts reflect this finding. It thus appears that the CBCA, in its current form, is limited in its effectiveness assessing deception in children. Perhaps supplementing the CBCA with children's ratings of clarity and confidence would improve the discriminability of the measure without the confounding effects of age and event familiarity.

ACKNOWLEDGMENTS

This article is based upon work supported by the National Science Foundation conducted by Kathy Pezdek under Grant No. SES-0111240. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

This research was conducted as Iris Blandon-Gitlin's Master of Arts thesis. Kathy Pezdek, in the Department of Psychology at Claremont Graduate University, can be contacted by e-mail at kathy.pezdek@cgu.edu.

REFERENCES

- Boychuk, T. D. (1991). *Criteria-based content analysis of children's statements about sexual abuse: A field-based validation study*. Unpublished doctoral dissertation, Arizona State University, Tempe.
- Buck, J. A., Warren, A. R., Betman, S. I., & Brigham (2002). Age differences in criteria-based content analyses scores in typical sexual abuse interviews. *Applied Developmental Psychology, 23*, 267–283.
- Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113*, 403–439.
- Craig, R. A., Scheibe, R., Raskin, D. C., Kircher, J. C., & Dodd, D. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. *Applied Developmental Science, 3*, 77–85.
- Kohnken, G., Schimossek, E., Aschermann, E., & Hofer, E. (1995). The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology, 80*, 671–684.
- Lamb, M. E., Sternberg, K. J., Esplin, P. W., Hershkowitz, I., Orbach, Y., & Hovav, M. (1997). Criteria-Based Content Analysis: A field validation study. *Child Abuse and Neglect, 21*, 255–264.
- Lamers-Winkelmann, F., & Buffing, F. (1996). Children's testimony in the Netherlands: A study of statement validity analysis. *Criminal Justice and Behavior, 23*, 304–321.
- Landry, K. L., & Brigham, J. C. (1992). The effect of training in criteria-content analysis on the ability to detect deception in adults. *Law and Human Behavior, 16*, 663–676.
- Pezdek, K., Finger, K., & Hodge, D. (1997). Planting false childhood memories: The role of event plausibility. *Psychological Science, 8*, 437–441.
- Pezdek, K., & Hodge, D. (1999). Planting childhood memories in children: The role of event plausibility. *Child Development, 70*, 887–895.
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., et al. (2004). Detecting deception in children: Event familiarity affects Criterion-Based Content Analysis ratings. *Journal of Applied Psychology, 89*, 119–126.
- Pezdek, K., & Taylor, J. (2000). Discriminating between accounts of true and false events. In D. F. Bjorklund (Ed.), *Research and theory in false-memory creation in children and adults*. Mahwah, NJ: Erlbaum.
- Raskin, D. C., & Esplin, P. W. (1991a). Statement Validity Assessment: Interview procedures and content analysis of children's statements of sexual abuse. *Behavioral Assessment, 12*, 265–291.
- Raskin, D. C., & Esplin, P. W. (1991b). Assessment of children's statements of sexual abuse. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 153–164). Washington, DC: American Psychological Association.
- Sedlak, A. J., & Broadhurst, D. D. (1996). *Third national incidence study of child abuse and neglect: Final report*. Washington, DC: US Department of Health and Human Resources, Administration for Children and Families.
- Steller, M. (1989). Recent developments in statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 135–154). Dordrecht, Netherlands: Kluwer.
- Tye, M. J. C., Amato, S. L., Honts, C. R., Devitt, M. K., & Peters, D. (1999). The willingness of children to lie and the assessment of credibility in an ecologically relevant laboratory setting. *Applied Developmental Science, 3*, 92–109.
- 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*(3), 261–283.
- Yuille, J. C. (1988). The systematic assessment of children's testimony. *Canadian Psychology, 29*, 247–262.
- Zaparniuk, J., Yuille, J. C., & Taylor, S. (1995). Assessing the credibility of true and false statements. *International Journal of Law Psychiatry, 18*, 343–352.