# Identifying Applicant Faking in Job Interviews: Examining the Role of Criterion-Based Content Analysis and Storytelling

Nicolas Roulin

*Saint Mary's University*

Deborah M. Powell

*University of Guelph*

Author note

Correspondence should be send to Nicolas Roulin, Department of Psychology, Saint Mary's University, 923 Robie Street, Halifax, NS, Canada, B3H 3C3, E-mail: nicolas.roulin@smu.ca, phone: +1 902 420 5831, fax: +1 902 496 8287.

This research was supported by an Insight Development Grant from the Social Sciences and Humanities Research Council of Canada (SSHRC Grant # 430-2015-00098).

The authors thank Bahareh Khosropanah, Brittany Young Tenn, Dale Camuyong, Manu Garekar, Zach Unrau, Sarah Sutey, Timothy De Pass, Brianne Gayfer, Kaytlin Fatovic, Megan Piccolo, Aaron Vieira, Todd Reesor, Katie Zieba, Emese Orosz, and Rachael Jones-Chick for their help with the data collection, transcription, and coding.

**Abstract**

Applicants' use of faking tactics could threaten the validity of employment interviews. We examined Criterion-Based Content Analysis (CBCA), an approach used in legal contexts, as a potential indicator of interviewee faking. We also examined the moderating role of storytelling in the faking-CBCA relationship. We conducted one experimental study, with 100 interviewees receiving instructions to respond honestly vs. to exaggerate/invent responses, and one mock interview study, with self-reported faking from 111 interviewees. Responses were recorded, transcribed, and coded for CBCA and storytelling. Faking was associated with CBCA when interviewees freely engaged in faking tactics, an overall CBCA indicator was used, and interviewees' responses contained story features. Additional analyses highlight that CBCA-based assessments of faking/honesty could reach up to 63.4% accuracy.

# Identifying Applicant Faking in Employment Interviews: Examining the Role of Criterion-Based Content Analysis and Storytelling

The employment interview is one of the most commonly used methods for assessing job candidates. Although the goal of the interview (for the organization) is to assess candidates' job-related qualifications, interview scores are influenced by more than a candidate's job-related skills and abilities. Job applicants may sometimes engage in faking, that is, choose to misrepresent their skills and abilities or past experiences (Levashina & Campion, 2007). Faking in the interview is especially problematic because interviewers are not particularly good at detecting it. For example, Roulin, Bangerter, and Levashina (2015) reported that only 12 to 19% of faking tactics used by applicants were accurately detected across multiple studies. If many applicants are using deceptive tactics in interviews, yet few interviewers can accurately detect such tactics, then honest candidates may be disadvantaged, and the validity of the interview may be affected (Levashina, Hartwell, Morgeson, & Campion, 2014). One reason for interviewers' inability to detect faking is that they rely on the wrong cues of deception. For instance, they may extensively rely on non-verbal behaviors to assess candidates' honesty, but those cues are generally invalid (Schneider, Powell, & Roulin, 2015).

An alternative approach to detecting faking could be to analyze the content and structure of interviewees' answers, an approach that has primarily been used as evidence in legal proceedings (Sporer, 1997; Vrij, Granhag, & Porter, 2010). The idea behind content-oriented approaches is that statements based on actual self-experienced events differ in a number of ways from fantasy or deceptive statements (Steller & Koehnken, 1989). For instance, research (e.g., Sporer, 1997) has highlighted Criterion-Based Content Analysis (CBCA), which includes criteria like logical structure or quantity of details as a valid approach for differentiating truthful from

untruthful accounts. Although a few content-oriented cues have been recently explored in the interview context (Culbertson, Weyhrauch, & Waples, 2016; Schneider et al., 2015), no research has systematically examined the CBCA system as an indicator of applicant faking. The CBCA approach has been traditionally used on the transcripts of accounts of events, or stories, in which participants were instructed to describe either a self-experienced or an invented event. In the interview context, CBCA could be an effective indicator of faking when interviewees are provided with the opportunity to describe a specific experience (when answering past behavior interview questions, Janz, 1982), and when their responses include the narrative elements of storytelling (Bangerter, Corvalan, & Cavin, 2014).

The present research systematically applies CBCA as a potential indicator of applicant faking in employment interviews, and examines if CBCA is more effective when applicants' responses include storytelling features. We propose to test this in two studies: one lab study with experimentally-manipulated faking levels and one mock interview study with self-reported applicant faking. Together, this research contributes to the personnel selection and applicant faking literatures in interviews in several ways. Theoretically, it examines the transferability of a coding system that has demonstrated promising results in legal proceedings to the personnel selection context. Practically, it represents the initial step towards developing a systematic method that organizations could apply to detect when applicants fake, for instance by training interviewers to identify valid cues to deception in interviewee's discourse (and taking away the distraction of non-verbal content) or by developing algorithms based on valid cues to automatically estimate interviewee's level of honesty.

**Applicant Faking in the Interview**

Applicant impression management (IM) in the employment interview has been defined as a conscious attempt to influence one's image during interactions (Ellis, West, Ryan, & DeShon, 2002). IM could be honest, for example, making sure that one's true skills and abilities are highlighted in an interview. However, IM can also be deceptive (Levashina & Campion, 2006, 2007). Faking (or deceptive IM) in the employment interview has been defined as "the conscious distortions of answers to the interview questions in order to obtain a better score on the interviewer and/or otherwise create favorable perceptions" (Levashina & Campion, 2007, p. 1639). Faking can include tactics close to lying, such as extensive image creation (inventing skills or borrowing experiences from a co-worker), but also milder tactics like slight image creation (exaggerating or embellishing), image protection (concealing information), and deceptive ingratiation (insincerely praising the interviewer).

Faking is a potential concern for organizations and interviewers, who want to know if applicants are honestly qualified or just pretending to be (Arthur, Glaze, Villado, & Taylor, 2010). It may thus introduce a source of inaccuracy in interviewers' assessments of applicants' qualifications (Levashina et al., 2014). Interviewers generally perceive themselves as capable of detecting when applicants use faking (Culbertson et al., 2016; Robie, Tuzinski, & Bly, 2006). Yet, both experimental and field studies suggest that interviewers actually perform quite poorly at faking detection (Roulin, Bangerter, & Levashina, 2014; Roulin et al., 2015). Such results are also line with past research on deception detection in general, suggesting that individuals rarely surpass chance level when attempting to identify lies in social situations (Bond & DePaulo, 2006, 2008).

Research in the legal context shows that deception detection improves when people use valid cues of deception instead of stereotypical cues (DePaulo et al., 2003; Vrij et al., 2010). For

instance, police officers' detection level is higher when they focus on story-related cues (e.g., vagueness, contradictions) instead of non-verbal cues (e.g., gaze aversions, fidgeting) (Mann, Vrij, & Bull, 2004). One potential way interviewers could be made more effective at detecting faking thus involves relying on valid (and ignoring less-valid) cues of faking (Roulin et al., 2015; Van Iddekinge, Raymark, & Roth, 2005). Yet, we must first identify which cues are associated with faking in job interviews.

Two recent studies have examined cues to applicant faking. Schneider et al. (2015) investigated 19 cues based on the literature on deception detection in general (DePaulo et al., 2003), and examined their relationships with applicants' self-reported use of the four types of faking tactics described earlier. They found very few valid non-verbal cues to faking, and those that were significantly associated with faking had small effect sizes. There seemed to be more promise for verbal cues like silences or speaking quickly. However, Schneider et al. (2015) did not specifically analyze the *content* of the interviewees' answers. Culbertson et al. (2016) compared the occurrence of 14 cues in mock interviews of individuals asked to lie vs. respond honestly. They found significant differences between honest and deceptive answers for 11 cues (e.g., deceptive answers were rated as more implausible), although effects were small to moderate in size. Interestingly, some of the cues used in that study belong to CBCA (Sporer, 1997), but the system was not used comprehensively. Moreover, they compared honest vs. deceptive responses, but did not examine different types of faking tactics (e.g., slight vs. extensive image creation). In this study we systematically apply the CBCA system to analyze the content and structure of interviewees' answers as a way to identify faking.

**Criterion-Based Content Analysis for Faking Detection**

Analyzing the content of written statements for cues to credibility has primarily been used as evidence in legal proceedings such as with eye witness testimonies (Vrij & Mann, 2006). The idea behind content-oriented approaches is that statements based on actual self-experienced events differ in a number of ways from statements not based on direct experiences (e.g., fantasy, borrowing from someone else), and such differences can be captured using specific cues or response styles (Steller & Koehnken, 1989). For instance, self-experienced events contain more qualitative details than invented events (Sporer & Sharman, 2006). Although there is a variety of coding systems measuring such differences in the witness testimony and deception detection literatures, CBCA has emerged as the most examined and valid technique (Sporer, 1997; Vrij, 2005).

CBCA originates from research on expert testimony (Sporer, 1982; Steller & Koehnken, 1989). The CBCA technique, as typically used in court settings, has three key steps. The first step is a semi-structured interview, during which the person recounts his or her story in a free recall style; this interview is audio-taped. Next, the interview is transcribed. Finally, the transcript is coded, by trained coders, for the CBCA criteria. The CBCA coding system comprises 19 criteria and coders judge the absence or presence of each criterion. There are both cognitive and motivational factors that influence CBCA scores (Vrij, 2005). From a cognitive perspective, the presence of more detailed and coherent descriptions of events (e.g., logical structure, quantity of details) in a statement is likely to indicate that it is truthful because such criteria are difficult to fabricate. From a motivational perspective, honest responders are generally less worried about looking credible than liars. As such, honest statement can include spontaneous corrections or admissions of lack of memory. The presence of each criterion strengthens the hypothesis that the account is based on genuine personal experience (Vrij &

Mann, 2006). Therefore, although CBCA criteria may have some value individually, an overall CBCA score is a better indicator of honesty.

The key strengths of the CBCA coding system are that meta-analytical findings suggest that (1) CBCA coding reaches adequate inter-rater agreement levels, and (2) it can help distinguish honest from deceptive statements above chance levels (Vrij, 2005). As an example, using 40 transcripts of truthful events and 40 transcripts of made-up events, Sporer (1997) found a multiple correlation of $R = .42$ between objective truth status and all CBCA criteria. Overall, trained CBCA coders can achieve up to 90% accuracy in deception detection (Griesel, Ternes, Schraml, Cooper, & Yuille, 2013; Vrij, 2005), much higher than the people's average detection scores (54%; Bond & DePaulo, 2006). As a result of these strengths, CBCA is now widely accepted as a valid method of credibility assessments in both European and North American courts (Griesel et al., 2013).

Although CBCA has been extensively used in forensic psychology research and applied in the legal context, it has not been used in assessing the veracity of applicants' responses in employment interviews. There are two weaknesses of CBCA coding, which could limit the applicability of this technique to employment interviews. First, the CBCA coding system was designed for use with statements provided in interviews in which free recall was encouraged, and prompting was kept at a minimum. Therefore, it may be limited to interviews that are structured (i.e., with limited probing) and include past-behavior oriented question (i.e., to elicit more free recall – but see next section). Second, the technique was designed to be used with written transcripts, rather than audio or video, and so its use may be limited to cases where interviews can be recorded and transcribed. Indeed, Vrij (2005) reports that CBCA experts are typically not in favor of assessing videotaped statements as watching a videotape might distract the CBCA

assessor from his or her assessment task. Despite these limitations, CBCA could be a potentially valid method to distinguish honest from faked responses, and ultimately identify applicants who fake in employment interviews, when the conditions of free recall and transcription can be met. We thus propose to examine the relationships between CBCA and faking. In this research, we focus on two types of faking tactics (i.e., slight and extensive image creation), which involve applicants embellishing or inventing qualifications or past work experiences, and are thus conceptually more similar to lies and more relevant to the CBCA system than other faking tactics (e.g., deceptive ingratiation). Based on the CBCA literature described above, we expect the following:

*Hypothesis 1*: Faking use will be negatively associated with CBCA scores.

**The Role of Storytelling in Faking Detection**

Past research using CBCA has focused on the transcripts of accounts of events, or stories, in which participants were instructed to describe either a self-experienced or an invented event. In the context of an employment interview, the process of describing an event would be similar to answering a past behavior interview question. Past behavior questions ask candidates to talk about past work experiences or life events, and are designed to elicit information about specific competencies (Janz, 1982). Bangerter et al. (2014) found that scores on responses to these questions are higher to the extent that applicants respond with stories, which they defined as "a set of events related to a unique past episode, characterized by a unity of time or action, which constituents often linked by temporal markers (p. 598)." Responses lacking such features (i.e., pseudo-stories or non-stories) obtained lower ratings by interviewers.

Bangerter et al. (2014) argued that interviewers are sensitive to the narrative content in applicants' responses, and that story-like answers are perceived as being more credible by

interviewers, who ultimately evaluate them more positively. We argue that storytelling (whether the response has the narrative elements of a story) could also be an important piece of the value of the CBCA system as a way to identify faking in interviews. CBCA is a method that assesses the veracity of a statement using characteristics like logical structure, level of details, or description of interactions (Vrij & Mann, 2006). In legal proceedings, this method works because eye witnesses' (or suspects') re-telling of events generally include these characteristics. In the interview context, the CBCA method may thus work better (or only) to identify faking if the discourse that is analyzed (i.e., the interviewee's response) has the appropriate characteristics of a story, as described by Bangerter et al. (2014). In contrast, responses that do not involve story features (e.g., in which interviewees simply list qualifications that they possess) may not include the type of content necessary for CBCA to work as a faking indicator. We therefore hypothesize that CBCA scores will be (negatively) associated with faking, particularly when the response to a past-behavior question is a "story". When the response is a "pseudo-story" or a non-story, the indicator would be less effective.

*Hypothesis 2:* Story structure will be a moderator of the relationship between faking use and CBCA scores, such that the relationship will be stronger for responses that are stories (vs. pseudo or non-stories).

## Overview of Studies

We examined the relationships between CBCA, story structure, and applicant faking in employment interviews and tested our two hypotheses in two studies. In Study 1 we used an experimental design. Interviewees were asked eight interview questions, with a randomization device used before each question instructing interviewees to respond in an honest, slightly deceptive, or extremely deceptive way. Such a design allows us to obtain a precise and

controlled measure of applicant faking use for each response, thus increasing the internal validity of our findings. However, such instructions limit our understanding of the effectiveness of the CBCA indicator in a more natural setting, with interviewees freely engaging in faking. In complement, Study 2 relied on interview data from Schneider et al. (2015), with mock interviews conducted by professional interviewers and interviewees' self-reports of faking for the whole interview. This second study allows us to test the value of the CBCA indicator when interviewees spontaneously use faking tactics, and thus increases the external validity of our findings. In both studies, the interviews were video-recorded, transcribed, and coded for CBCA and story production by independent groups of trained coders blind to level of faking used.

## Study 1

### Methods

#### Sample and Procedure

We recruited 100 students from a Canadian business school, with 53 MBA students (34% female, 38% of Caucasians) and 47 Bachelor of Commerce students (51% female, 60% of Caucasians). They participated in a structured mock interview, in exchange for a $15 gift card or course credit. To increase the realism of the interviews, participants were asked to dress professionally and to imagine that they were interviewing for a general management position. All interviews were conducted by the same trained research assistant, following a strict protocol. They included eight behavior description questions, each measuring a specific competency (e.g., teamwork, communication). To increase the likelihood that responses contained enough story-like content, interviewees were encouraged to structure their responses using the STAR method (i.e., describing the situation, tasks, actions, and results; Bangerter et al., 2014). The interviewer

was instructed to repeat the question or use a standardized probe ("would you like to add something?") only when the response provided by the interviewee was very short or incomplete.

To elicit responses with various faking levels, we designed a randomization system. Before the interviewer asked each question, interviewees were instructed to roll a dice and to take note of the result. This was done behind a small screen installed on the interview table, so that the interviewer was blind to the dice outcome. When the dice showed "1" or "2", interviewees were instructed to use an honest response strategy (i.e., a fully accurate description of a previous experience highlighting their skills). When it showed "3" or "4", they had to use slight image creation (i.e., including slight exaggerations, embellishments, or transformations of their experience). Finally, when it showed "5" or "6", they had to use extensive image creation (i.e., including extreme forms of exaggerations or even inventions of events). Participants were provided with definitions and examples for the three strategies derived from previous faking research (Levashina & Campion, 2007; Roulin, 2016) prior to the first question. Overall, our randomization system led to 794 responses: 283 honest, 276 slightly deceptive, and 235 extensively deceptive (and in 6 cases the interviewee was unable to answer the question). We also note that the interviewer rated the quality of each answer provided by the applicant using 5-level behaviorally-anchored rating scales designed to assess each of the eight competencies.

**Coding**

*Criterion-Based Content Analysis (CBCA) coding.* Interviewees' answers were transcribed by a trained research assistant, leading to 794 (i.e., 100 x 8, minus six unanswered questions) transcribed responses. Responses were then independently coded by two trained coders, who scored each transcribed response on a set of 14 CBCA indicators derived from past research (Sporer, 1997; Vrij, 2005; Vrij & Mann, 2006). We kept most of the original indicators

used in witness testimony research, but excluded indicators not applicable to a job interview context (e.g., about the crime perpetrator). The complete list of CBCA indicators, definitions, and examples can be found in Table 1. Each indicator was coded on a 0-2 scale, with 0=absent from the response", "1=some indication of presence in the response", and "2= clear/strong indication of presence in the response". The two coders received extensive training on CBCA coding by one of the authors, following the approach described in past research (e.g., Vrij, Mann, Kristen, & Fisher, 2007). The training included a two-hour session providing information about the theoretical foundations of CBCA, as well as definitions and examples for each indicator. One author and the two coders then coded several response transcripts, compared and discussed their results, and clarified potential misunderstandings. The two coders then independently coded the transcripts of ten interviews, inter-coder agreement was verified, and example disagreements were discussed. Finally, the coders independently coded all the remaining responses. The overall inter-rater consistency was good (*ICC* = .76), so the CBCA scores of the two coders for each response were aggregated. We computed a total CBCA score by summing the scores on the 14 indicators.

*Story coding.* Another group of two trained coders coded transcripts for story production, using the coding system developed by Bangerter et al. (2014). Each response was coded 1 if it was a "story" (i.e., a description of a unique situation, with contextual information, a set of unique events and actions, and statements linked with temporal cohesion markers). It was coded 0 if it was a "pseudo-story" (i.e., a description of a general context including a number of situations or general statements, but not a description of a specific situation, or set of actions) or a "non-story" (i.e., a response that does not fit with any of the two previous descriptions). Inter-

rater consistency was good (*ICC* = .79), so the story scores of the two coders for each response were aggregated for each answer.

**Results**

All analyses were performed at the response level. Results are presented in Table 2. We first examined potential differences in the CBCA overall scores between the three faking levels using an ANOVA. Contrary to our first hypothesis that CBCA scores would be lower for faked responses, results showed no significant difference in overall scores between responses in the honest, slightly deceptive, and extensively deceptive conditions, $F(2, 794) = .182$, $p = .979$. We further explored potential differences at the indicator level with a MANOVA. We found no overall effect, $F(26, 1550) = 1.273$, $p = .162$, $\eta_p^2 = .017$. Yet, subsequent indicator-level ANOVAs highlighted a small but significant difference for the *Unexpected Complications* indicator, $F(2, 794) = 3.031$, $p < .05$, $\eta_p^2 = .008$, with higher scores for more deceptive responses (again, contrary to Hypothesis 1).

We examined the moderating role of story production in the relationship between applicant faking and CBCA scores with multiple regressions. Story and faking levels were entered as predictors of CBCA in Step 1, and the interaction was added in Step 2. Results can be found in Table 3. They highlight significant main effects of Story (but not faking level) on CBCA in Step 1. In other words, story-like responses scored higher on CBCA. In addition, and contrary to our second hypothesis, we did not find any significant Story x Faking interaction in Step 2.

We also note that the performance ratings were significantly correlated with the overall CBCA score ($r = .30$, $p < .01$), but they were unrelated to the faking manipulation ($r = -.06$, $p = .10$) and unrelated to the story level ($r = .01$, $p = .76$).

## Study 2

### Methods

#### Sample and Procedure

We used the interview data from Schneider et al. (2015). A total of 111 undergraduate students in business or psychology from a different Canadian University (70% female; 70% Caucasians; mean age = 19.8) participated in videotaped mock interviews with one of four management consultants. Each interview consisted of three past behavior description and three situational questions. Because CBCA is only relevant when rating statements based on past situations/actions, we only used the answers to the three past behavior description questions, which measured communication, organization, and time management.

#### Measures and Coding

*Faking use*. After the interview, each interviewee completed a 14-item faking measure from Levashina and Campion's (2007) Interview Faking Behavior Scale. For the present study, we focused on interviewees' self-reports of slight image creation (4-item scale, $\alpha$ = .82) and extensive image creation (4-item scale, $\alpha$ = .73). We note that, although we focus only on the content of the behavior description questions in this study, interviewees completed the IM measures about the complete interview.

*CBCA coding.* Interviewees' answers to the three behavior description questions were transcribed by two trained research assistants, leading to 333 (111 x 3) transcribed responses. Responses were then independently coded for CBCA, using the same coding procedure as in Study 1. A group of two trained coders (one being the same as in the first study) coded all transcribed responses using the 14 CBCA indicators. Inter-rater consistency was good (*ICC* = .73), so the CBCA scores of the two coders for each response were aggregated. We then

computed the average (0-2) scores for each CBCA indicator across the three questions, and finally computed an overall CBCA score (0-28) for each interviewee.

*Story coding.* Responses were independently coded for story production by another group of four trained coders (the same two as in Study 1, plus two new coders), using the same coding procedure as in Study 1. To obtain a story score at the interviewee level, we computed the average story score (i.e., 0-1) across the three behavior description questions for each coder. Inter-rater consistency was excellent (*ICC* = .88), so the story scores of the four coders for each interviewee were aggregated.

*Interview performance.* Interviewers rated the quality of each answer to the three behavioral questions on 1-5 behaviorally-anchored rating scales, and we computed an average performance score at the interview level.

**Results**

We first examined the direct relationships between our indicators of applicant faking (i.e., CBCA) and self-reported use of faking (i.e., slight and extensive image creation) by interviewees. Because the faking data were measured at the interview level, all analyses were performed at the interview level. Table 4 describes the correlations for the overall CBCA score, as well as for each CBCA indicators. In line with Hypothesis 1, the general CBCA score was negatively associated with both slight ($r$ = -.24, $p$ < .01) and extensive image creation ($r$ = -.20, $p$ < .05). Moreover, we did not find any significant relationship between faking and any individual CBCA indicator.

We examined the moderating role of story production in the relationship between applicant faking and CBCA with multiple regressions. Story and slight/extensive image creation were entered as predictors of CBCA in Step 1, and the interaction was added in Step 2. Results

can be found in Table 5. They highlight significant main effects of both story and faking (both with slight and extensive image creation) on CBCA in Step 1. Importantly, we found a significant story x faking interaction in Step 2 for both slight image creation, $b$ = -2.56, $SE$ = 1.00, $p < .01$ and extensive image creation, $b$ = -2.80, $SE$ = 1.11, $p < .05$. The moderating role of story production is illustrated in Figure 1. In line with Hypothesis 2, the negative relationship between faking and CBCA was only visible when interviewees produced more story-like responses.

We also note that the interview performance ratings were significantly correlated with the overall CBCA score ($r$ = .22, $p < .05$), but it was unrelated to faking use ($r$ = -.00, $p$ = .99 for slight image creation and $r$ = .12, $p$ = .23 for extensive image creation) and only moderately related to the story level ($r$ = .19, $p$ = .05).

## Discussion

### Contribution to Interview Faking Research

Understanding and detecting applicant faking has been identified as both an important direction for future employment interview research (Levashina et al., 2014) and a practical issue for organizations (Arthur et al., 2010). Recent efforts to identify valid cues to applicant faking have found limited support for non-verbal cues, but more promising results for verbal content (Culbertson et al., 2016; Schneider et al., 2015). Building on those encouraging earlier findings, the present research represented an initial attempt to employ a systematic content-based approach based on CBCA, a method that has been shown to effectively distinguish honest from deceptive statements in the legal context (e.g., Sporer, 1997; Vrij, Akehurst, Soukara, & Bull, 2002; Vrij et al., 2007), in the employment interview context. Overall, our results suggest that CBCA can potentially be a valid indicator of faking, but highlights boundary conditions for CBCA validity.

More precisely, our results suggest that interviewee faking and CBCA are associated only when three conditions are present: (a) interviewees are free to engage in faking tactics, (b) an overall CBCA indicator is used, and (c) interviewees' responses contain story features.

The first condition implies that the exaggerated or invented statements are more accurately captured by CBCA coding when interviewees are responding in a natural interview setting. Indeed, we did not find any evidence for CBCA being associated with faking when interviewees were instructed to respond honestly vs. deceptively (in Study 1), but we did find evidence when interviewees could spontaneously decide to engage in faking or not (in Study 2). This may be explained by the motivational reason for CBCA effectiveness. Deception is captured by CBCA indicators partly because liars try harder to control their speech to appear credible, thus producing less detailed responses and admitting less errors our doubts about their statements (Vrij & Mann, 2006). This motivational factor may be especially present in a realistic interview situation, in which interviewees are particularly concerned with credibility, and adapt their faking tactics to appear credible. In contrast, such concerns may be reduced in a lab experiment, leading interviewees to engage in less controlled responses when instructed to fake. In the legal context, this motivational argument has been similarly used to explain why CBCA indicators tend to be less effective in lab than in field settings (Vrij, 2005). It is also possible that the instructions provided in Study 1 restricted interviewees' ability to use faking tactics freely and more effectively. For instance, in a real selection context interviewees engage in faking particularly when they have to demonstrate possessing a job-related skill or ability that they perceive to be lacking (Marcus, 2009). It may be that our randomization system led to situations where the interviewees were instructed to fake on questions about skills that they did actually possess, but to be honest on questions about skills that they were lacking.

The second condition involves using overall scores. In our second study, the faking-CBCA relationship was observed only when using the overall CBCA score, but not with individual CBCA indicators. This suggests that interviewers and organizations should refrain from focusing on a few specific cues (e.g., logical structure or level of details), but should rather assess interviewees' honesty using a general score based on a variety of CBCA indicators. This result is aligned with the general practice with real-life legal cases (Vrij et al., 2007).

The third condition involves interviewees providing story-like responses. The finding that CBCA is associated with faking only when interviewees' responses featured the narrative elements of a story, but not when they lacked such features, is consistent with past research on both job interviews and CBCA. More precisely, interview research suggests that responses to past-behavioral questions are evaluated more positively by interviewers when they feature the narrative elements of a story, such as detailed descriptions of the situations, tasks, actions, or results (Bangerter et al., 2014). The positive effect of story features is likely because story-like answers are seen as more credible. Similarly, CBCA research suggests that CBCA indicators more effectively distinguish honest from deceptive statements when the interview is designed to elicit verbal cues to credibility (Griesel et al., 2013; Vrij et al., 2007). As a practical implication, this result implies that organizations interested in using CBCA to identify faking should start by designing interviews that elicit or facilitate interviewees' storytelling (e.g., using the STAR technique).

Interestingly, our data also highlight that CBCA was positively related to ratings of interview performance in both studies. This suggests that applicants whose answers contained more CBCA indicators (which, according to CBCA research, should signal more truthful statements) were also evaluated more positively by interviewers.

**Potential Practical Implications**

Is CBCA practically effective to detect faking? As an illustration of its potential, we explored its effectiveness for detecting interviewees who engaged in extensive image creation with different CBCA cutoff scores, using data from Study 2. More precisely, 59 interviewees can be considered as fakers (i.e., reported engaging in extensive image creation), whereas 42 can be considered as honest (i.e., did not engage in extensive image creation). We computed the percentage of interviewees correctly identified as fakers, correctly identified as honest, and the overall detection accuracy, when using various CBCA cutoff scores (Figure 2 - CBCA score theoretically range from 0 to 28, with a mean of 7.42 in this study). With low cutoff scores (e.g., all interviewees scoring higher than 4 are labelled as honest), most of the honest interviewees are correctly identified, but most of the deceptive interviewees are not. With high cutoff scores (e.g., 11), most deceptive interviewees are correctly identified, but most honest ones are not. We obtained the highest detection rate (honest and deceptive combined) for a cutoff score of 9, with a correct detection of 63.4% of interviewees. The results were similar when using only interviewees using story features in their responses (i.e., 70 individuals, 42 being deceptive, and 65.7% correct detection).

Such correct detection scores are much higher than detection scores obtained with people trying to detect faking tactics (55%; Roulin, 2016) or deception (54%; Bond & DePaulo, 2006), or simply chance level (50%). However, they are on the lower end of those obtained with CBCA in the legal context (55-90%; Vrij, 2005). This is possibly because courts and job interviews are arguably very different contexts. For instance, some CBCA dimensions (e.g., spontaneous corrections, self-depreciation) may be more common and thus very useful to identify deception in courts, but are rarely present in interviews and thus less valid for detecting faking. Although

we believe that CBCA shows potential to identify fakers, we also want to emphasize that the present research is only the initial step in that direction, and that more research is needed before we can recommend its application to organizations. For instance, it is important to find CBCA cutoff levels that identify most fakers without incorrectly labelling honest respondents. In our illustration, the cutoff score of 9 allowed us to correctly identify over 90% of fakers but 80% of the honest candidates were incorrectly labelled as a faker.

**Limitations and Future Research Directions**

This research has limitations that create opportunities for further examination of indicators of faking in the future. First, our CBCA coders were initially complete novices in this technique and were only provided with a few hours of training. Although our high interrater agreement scores provide some evidence of training effectiveness, our coders were not CBCA experts. This increases the practical relevance of our findings, because organizations interested in using this technique are likely to provide a similar level of training to managers. Yet, future research may also examine if CBCA is a more effective indicator of faking when true experts are scoring transcripts. Similarly, research could explore the effectiveness of training managers or interviewers to use the CBCA technique.

Second, the interviewees involved in both studies were university students, with limited work and interviewing experience (except for the 53 MBA students involved in Study 1). Future research could thus try to replicate our results with more experienced interviewees, who may have more (or more detailed) job experiences to describe in their responses (Bangerter et al., 2014), thus making the CBCA indicator more effective to identify faking.

Third, a limitation of Study 2 is the relatively imprecise measure of faking used (i.e., self-reported use across the entire interview while the coding was done only on half of the questions). It is therefore possible that the self-reported faking scores may over-(or under) estimate the actual level of faking interviewees engaged in during the three past-behavioral questions.

Fourth, although self-reported measures of faking have been extensively used in the interview literature in the last decade (e.g., Levashina & Campion, 2007; Roulin et al., 2014) and interviewees in our second study were not applying for a real job, there is always a possibility that some interviewees' under-reported their use of faking. Future studies may attempt to use other measures of interview faking, for instance based on new techniques derived from the bogus item method (Levashina, Morgeson, & Campion, 2009).

Finally, the CBCA system could be fine-tuned to be more effective. For instance, some CBCA indicators (e.g., superfluous details, self-depreciation) had very low mean scores in our studies, but also in past CBCA research. Future studies could consider leaving those categories out of the coding system to make it more efficient. Moreover, we acknowledge that transcribing interviews and coding for CBCA indicators could be too labor intensive in practice. As such, future studies could also explore the effectiveness of the CBCA methodology with purely qualitative (or clinical) judgments of trained interviews (Griesel et al., 2013). Alternatively, research could examine the effectiveness of automatic coding of interview recordings (Vrij et al., 2007). For instance, previous research in forensic psychology (e.g., Lee, Lusk, & Halperin, 2014) has used the Linguistic Inquiry Word Count Program and scoring categories derived from CBCA to examine fraudulent financial reports. Although Lee et al.'s categories might not directly be applicable to interview transcripts, similar approaches could be explored. More advanced technologies, such as machine learning or deep learning, could also be examined as

automatic ways to score transcripts or directly audio- or video-recordings. Research could also directly compare the accuracy of human raters trying to identify fakers to CBCA-based identification using the same video material.

## References

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*, 1086-1120. doi: 10.1016/j.leaqua.2010.10.010

Arthur, W., Glaze, R. M., Villado, A. J., & Taylor, J. E. (2010). The magnitude and extent of cheating and response distortion effects on unproctored internet-based tests of cognitive ability and personality. *International Journal of Selection and Assessment, 18*, 1-16. doi: 10.1111/j.1468-2389.2010.00476.x

Bangerter, A., Corvalan, P., & Cavin, C. (2014). Storytelling in the selection interview? How applicants respond to past behavior questions. *Journal of Business and Psychology, 29*, 593-604. doi: 10.1007/s10869-014-9350-0

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214-234. doi: 10.1207/s15327957pspr1003_2

Bond, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*, 477-492. doi: 10.1037/0033-2909.134.4.477

Culbertson, S. S., Weyhrauch, W. S., & Waples, C. J. (2016). Behavioral Cues as Indicators of Deception in Structured Employment Interviews. *International Journal of Selection and Assessment, 24*, 119-131. doi: 10.1111/ijsa.12135

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

Ellis, A. P. J., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology, 87*, 1200-1208. doi: 10.1037/0021-9010.87.6.1200

Griesel, D., Ternes, M., Schraml, D., Cooper, B. S., & Yuille, J. C. (2013). The ABC's of CBCA: Verbal Credibility Assessment in Practice. In S. B. Cooper, D. Griesel & M.

Ternes (Eds.), *Applied Issues in Investigative Interviewing, Eyewitness Memory, and Credibility Assessment* (pp. 293-323). New York, NY: Springer New York.

Janz, T. (1982). Initial comparisons of patterned behavior description interviews versus unstructured interviews. *Journal of Applied Psychology, 67*, 577-580. doi: 10.1037/0021-9010.67.5.577

Lee, C.-H., Lusk, E. J., & Halperin, M. (2014). Content Analysis for Detection of Reporting Irregularities: Evidence from Restatements during the SOX Era. *Journal of Forensic and Investigative Accounting, 6*, 99-122.

Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment, 14*, 299-316. doi: 10.1111/j.1468-2389.2006.00353.x

Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: Development and validation of an interview faking behavior scale. *Journal of Applied Psychology, 92*, 1638-1656. doi: 10.1037/0021-9010.92.6.1638

Levashina, J., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology, 67*, 241-293. doi: 10.1111/peps.12052

Levashina, J., Morgeson, F. P., & Campion, M. A. (2009). They don't do It often, but they do it well: Exploring the relationship between applicant mental abilities and faking. *International Journal of Selection and Assessment, 17*, 271-281. doi: 10.1111/j.1468-2389.2009.00469.x

Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology, 89*, 137-149. doi: 10.1037/0021-9010.89.1.137

Marcus, B. (2009). 'Faking' from the applicant's perspective: A theory of self-presentation in personnel selection settings. *International Journal of Selection and Assessment, 17*, 417-430. doi: 10.1111/j.1468-2389.2009.00483.x

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica, 46,* 69-85. doi: 10.2307/1913646

Robie, C., Tuzinski, K. A., & Bly, P. R. (2006). A survey of assessor beliefs and practices related to faking. *Journal of Managerial Psychology, 21,* 669-681. doi: 10.1108/02683940610690204

Roulin, N. (2016). Individual differences predicting impression management detection in job interviews. *Personnel Assessment and Decisions, 2,* 1-11.

Roulin, N., Bangerter, A., & Levashina, J. (2014). Interviewers' perceptions of impression management in employment interviews. *Journal of Managerial Psychology, 29,* 141-163. doi: 10.1108/JMP-10-2012-0295

Roulin, N., Bangerter, A., & Levashina, J. (2015). Honest and deceptive impression management in the employment interview: Can it be detected and how does it impact evaluations? *Personnel Psychology, 68,* 395-444. doi: 10.1111/peps.12079

Schneider, L., Powell, D. M., & Roulin, N. (2015). Cues to deception in the employment interview. *International Journal of Selection and Assessment, 23,* 182-190. doi: 10.1111/ijsa.12106

Sporer, S. L. (1982). A brief history of the psychology of testimony. *Current Psychological Reviews, 2,* 323-339. doi: 10.1007/bf02684465

Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology, 11,* 373-397. doi: 10.1002/(SICI)1099-0720(199710)11:5<373::AID-ACP461>3.0.CO;2-0

Sporer, S. L., & Sharman, S. J. (2006). Should I believe this? Reality monitoring of accounts of self-experienced and invented recent and distant autobiographical events. *Applied Cognitive Psychology, 20,* 837-854. doi: 10.1002/acp.1234

Steller, M., & Koehnken, 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.

Van Iddekinge, C. H., Raymark, P. H., & Roth, P. L. (2005). Assessing personality with a structured employment interview: Construct-related validity and susceptibility to response inflation. *Journal of Applied Psychology, 90*, 536-552. doi: 10.1037/0021-9010.90.3.536

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., 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., & 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., & Mann, S. (2006). Criteria-Based Content Analysis: An empirical test of its underlying processes. *Psychology, Crime & Law, 12*, 337-349. doi: 10.1080/10683160500129007

Vrij, A., Mann, S., Kristen, S., & Fisher, R. P. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*, 499-518. doi: 10.1007/s 10979-006-9066

**Table 1. CBCA Indicators, Definitions, and Examples**

| Indicator | Definition | Example |
|---|---|---|
| Logical structure | Statement is coherent and consistent, with different details that all describe the same course of events | "I was in charge of managing a team of 5 people. Those five individuals were hired by my predecessor, so I did not chose them. Yet, I had to lead them to complete this project on time., I thus assigned each subordinate a specific task to do" |
| Unstructured production | Information is NOT provided in a chronological time sequence, and expressive style showing lack of constraint and producing an unstructured presentation | "One time I had conflicting demands at work was when I had to manage Project A and Project B both at the same time. So I was at company XYZ and Project B was very important for my boss and time sensitive. But before that I started Project A that was my own initiative. So..." |
| Quantity of details | Statement involves a considerable amount of details | "It was in the summer of 2013 and I was not taking any course at that time. So I worked as the assistant manager in store XYZ for three months. The store was open 7 days a week and there were 12 employees working there. My role was to supervise them and..." |
| Contextual embedding | Reference to time and space as a basis for the story | "I started working in Store XYZ downtown in June 2014" |
| Descriptions of interactions | Descriptions of any kind of interaction | "I approached the customer, smiled at her, and welcomed her to the store" |
| Reproduction of conversations | Virtual replication of actual wording of some part of conversation | "I went to my coworker and then I asked him "*why didn't check the inventory yesterday"* and he said "*I was too busy…*" |
| Unexpected complications | Description of elements that were somewhat unexpected (e.g., unforeseen interruption or difficulty) | "I remember that there was a problem with the cash registry. It just did not open properly." |
| Unusual details | Emphasis on details that are uncommon but meaningful to the story (with a degree of concreteness and vividness) | "The client was old and seemed to have hearing difficulties. I noticed that she had some kind of hearing aid system around her ear. So I tried to speak slowly but loudly..." |
| Superfluous details | Details that are not essential to the main story | "I was talking to that customer about a TV he wanted to buy. I remember that he was wearing a blue football jersey. And I..." |
| Subjective mental state | Reports of feelings, thoughts, or cognitions experienced at the time of the event | "It was the first time I was making a presentation in front of the management team. I was quite stressed and I was worried I would make a mistake…" |
| Spontaneous corrections | Correcting oneself during the interview without prompting from the interviewer | "So I was working the company for about three months… no wait, I think it was 4 months already." |
| Lack of memory | Expressing concern that some parts of the statement might be incorrect | ‘‘I think’’, ‘‘Maybe’’, ‘‘I am not sure’’, etc. |
| Doubts about testimony | Anticipated objections against the veracity of one’s own description | ‘‘I know all this sounds really odd… but this is how this client reacted that day’’ |
| Self-deprecation | Mention of personally unfavorable or self-incriminating details | ‘‘Obviously it was stupid of me to forget that document because I knew it was important for the meeting. But…" |

**Table 2. Storytelling and CBCA Indicators by Faking Condition (Study 1)**

| | Honest *Mean* (*SD*) | Slight image creation *Mean* (*SD*) | Extensive image creation *Mean* (*SD*) | *ANOVA* *F-value* | $\eta_p^2$ |
|---|---|---|---|---|---|
| Story | 0.74 (0.40) | 0.80 (0.36) | 0.83 (0.37) | 3.650* | .009 |
| CBCA Total | 7.42 (2.22) | 7.31 (2.16) | 7.35 (2.36) | .182 | .000 |
| CBCA Logical structure | 1.84 (0.32) | 1.87 (0.27) | 1.85 (0.34) | .618 | .002 |
| CBCA Unstructured production | 0.25 (0.41) | 0.24 (0.40) | 0.23 (0.38) | .243 | .001 |
| CBCA Quantity details | 1.31 (0.61) | 1.25 (0.56) | 1.28 (0.57) | .769 | .002 |
| CBCA Contextual embedding | 1.11 (0.54) | 1.07 (0.52) | 1.07 (0.49) | .463 | .001 |
| CBCA Descriptions interactions | 0.86 (0.46) | 0.89 (0.46) | 0.87 (0.45) | .316 | .001 |
| CBCA Reproduction conversation | 0.15 (0.30) | 0.19 (0.36) | 0.18 (0.34) | .978 | .002 |
| CBCA Unexpected complications | 0.53 (0.45) | 0.60 (0.42) | 0.62 (0.44) | 3.031* | .008 |
| CBCA Unusual details | 0.60 (0.42) | 0.58 (0.41) | 0.59 (0.42) | .219 | .001 |
| CBCA Superfluous details | 0.19 (0.33) | 0.16 (0.32) | 0.15 (0.28) | .980 | .002 |
| CBCA Subjective mental state | 0.40 (0.40) | 0.37 (0.37) | 0.38 (0.38) | .639 | .002 |
| CBCA Spontaneous corrections | 0.05 (0.19) | 0.06 (0.17) | 0.06 (0.21) | .180 | .000 |
| CBCA Lack memory | 0.10 (0.28) | 0.09 (0.29) | 0.07 (0.21) | .782 | .002 |
| CBCA Doubts testimony | 0.00 (0.00) | 0.00 (0.04) | 0.00 (0.00) | 1.887 | .005 |
| CBCA Self Deprecation | 0.03 (0.17) | 0.03 (0.14) | 0.02 (0.11) | .685 | .002 |

*Note:* $N$ = 283, 276, and 235 for the three faking levels respectively. Story = Average scores of two coders, with ICC = .79; CBCA = Average scores of two coders, with ICC = .76; * $p < .05$, ** $p < .01$

**Table 3. Regressions Predicting CBCA Scores (Study 1)**

| | *Step 1* | | *Step 2* | |
|---|---|---|---|---|
| | *b (SE)* | *Beta* | *b (SE)* | *Beta* |
| Constant | 6.48** (.24) | | 6.39 (.42) | |
| Story | 1.35** (.20) | .23 | 1.46** (.50) | .25 |
| Faking level | -.09 (.09) | -.03 | -.04 (.22) | -.02 |
| Story x Faking level | | | -.06 (.25) | -.03 |
| *F* | | 22.68** | | 15.11** |
| $R^2$ | | .23 | | .23 |

*Note: N* = 800; Story = Average scores of two coders, with ICC = .79; CBCA = Average scores of two coders, with ICC = .76; * $p < .05$, ** $p < .01$. In order to capture the unobserved heterogeneity in CBCA scores at the interviewee level (because of the clustered nature of our data), we thus also conducted regressions including the cluster means of interviewees' CBCA in the estimated model following the procedure developed by Mundlak (1978, see also Antonakis, Bendahan, Jacquart, & Lalive, 2010). Results of these analyses mirrored the regressions results presented above.

**Table 4. Descriptive Statistics and Correlations with CBCA Indicators (Study 2)**

| | Descriptive Statistics | | | Correlations | | |
|---|---|---|---|---|---|---|
| | *Mean* | *SD* | *N* | *Story* | *Slight image creation* | *Extensive image creation* |
| Story | 0.61 | 0.27 | 111 | | | |
| Slight image creation | 1.94 | 0.76 | 108 | -.041 | | |
| Extensive image creation | 1.41 | 0.60 | 109 | -.057 | .576** | |
| CBCA Total | 6.87 | 2.15 | 107 | .416** | -.243** | -.200* |
| CBCA Logical structure | 1.76 | 0.25 | 109 | .304** | -.056 | -.095 |
| CBCA Unstructured production | 0.43 | 0.33 | 109 | .332** | .013 | -.009 |
| CBCA Quantity details | 0.88 | 0.51 | 109 | .508** | -.075 | -.017 |
| CBCA Contextual embedding | 1.22 | 0.41 | 109 | .514** | -.135 | -.076 |
| CBCA Descriptions interactions | 0.97 | 0.35 | 109 | .333** | -.042 | .050 |
| CBCA Reproduction conversation | 0.17 | 0.30 | 109 | .212* | .102 | .171 |
| CBCA Unexpected complications | 0.56 | 0.26 | 109 | .400** | .138 | .159 |
| CBCA Unusual details | 0.55 | 0.36 | 109 | .346** | -.049 | .018 |
| CBCA Superfluous details | 0.07 | 0.18 | 109 | .245* | -.085 | -.061 |
| CBCA Subjective mental state | 0.28 | 0.33 | 109 | .115 | .120 | .003 |
| CBCA Spontaneous corrections | 0.09 | 0.18 | 109 | .339** | -.021 | .068 |
| CBCA Lack memory | 0.28 | 0.32 | 109 | .202* | -.118 | -.176 |
| CBCA Doubts testimony | 0.01 | 0.05 | 109 | .090 | -.113 | -.013 |
| CBCA Self Deprecation | 0.02 | 0.09 | 109 | -.065 | .109 | -.082 |

*Note:* Story = Average scores of four coders, with ICC = .88; CBCA = Average scores of two coders, with ICC = .73; * $p < .05$, ** $p < .01$

**Table 5. Regression Predicting CBCA Scores (Study 2)**

| | Slight image creation | | | | Extensive image creation | | | |
|---|---|---|---|---|---|---|---|---|
| | *Step 1* | | *Step 2* | | *Step 1* | | *Step 2* | |
| | *b (SE)* | *Beta* | *b (SE)* | *Beta* | *b (SE)* | *Beta* | *b (SE)* | *Beta* |
| Constant | 5.76** (.71) | | 2.46 (1.68) | | 5.46** (.71) | | 2.69* (1.30) | |
| Story | 3.75** (.77) | .43 | 8.77** (2.09) | 1.01 | 3.72** (.78) | .43 | 7.74** (1.77) | .89 |
| Faking use | -.67** (.25) | -.24 | 1.02 (.70) | .36 | -.70* (.33) | -.19 | 1.23 (.83) | .33 |
| Story x Faking use | | | -2.56** (1.00) | -.86 | | | -2.80* (1.11) | -.72 |
| *F* | | 15.71** | | 13.27** | | 14.08** | | 11.99** |
| $R^2$ | | .25 | | .29 | | .22 | | .27 |

*Note: N* = 111; Story = Average scores of four coders, with ICC = .88; CBCA = Average scores of two coders, with ICC = .73; * $p < .05$, ** $p < .01$.

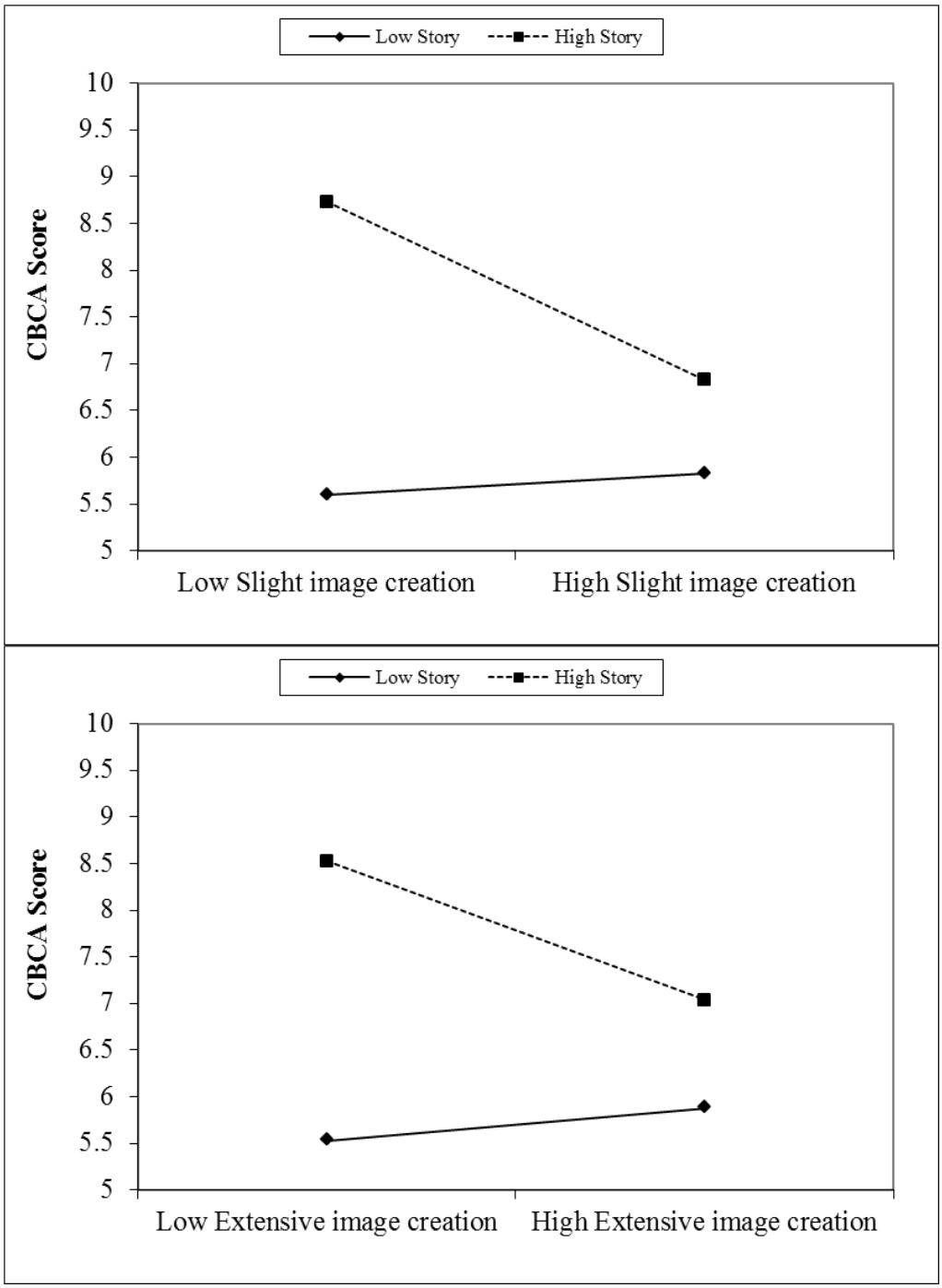


***Figure 1.*** CBCA score as a Function of Story and Slight/Extensive Image Creation (Study 2)

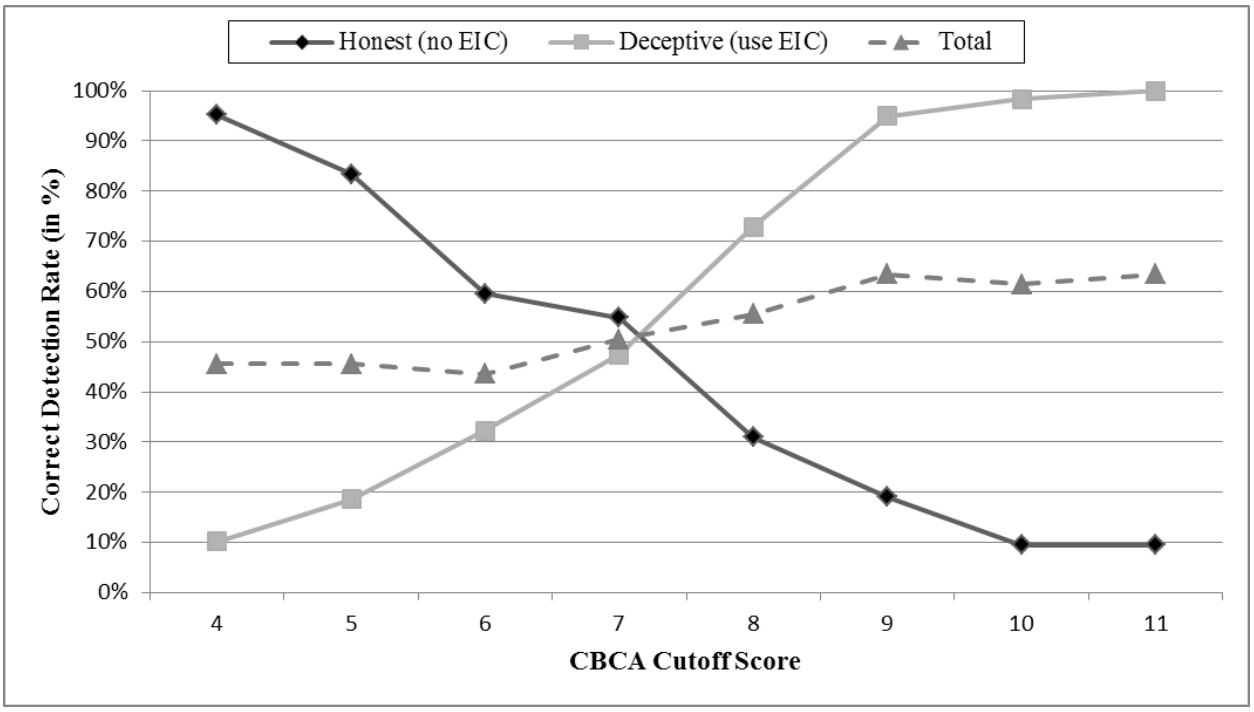


***Figure 2.*** Correct Detection of Extensive Image Creation Depending on CBCA Cutoff (Study 2)