Criteria-Based Content Analysis: An empirical test of its underlying processes

ALDERT VRIJ* & SAMANTHA MANN
Psychology Department, University of Portsmouth, Portsmouth, UK

Abstract
Criteria-Based Content Analysis (CBCA) is a tool to assess the veracity of written statements, and is used as evidence in criminal courts in several countries in the world. CBCA scores are expected to be higher for truth tellers than for liars. The underlying assumption of CBCA is that (i) lying is cognitively more difficult than truth telling, and (ii) that liars are more concerned with the impression they make on others than truth tellers. However, these assumptions have not been tested to date. In the present experiment 80 participants (undergraduate students) lied or told the truth about an event. Afterwards, they completed a questionnaire measuring “cognitive load” and “tendency to control speech”. The interviews were transcribed and coded by trained CBCA raters. In agreement with CBCA assumptions, (i) truth tellers obtained higher scores than liars, (ii) liars experienced more cognitive load than truth tellers, and (iii) liars tried harder to control their speech. However, cognitive load and speech control were not correlated with CBCA scores in the predicted way.

Keywords: deception, cognitive load, cbca

Introduction
Statement Validity Assessment (SVA) is a tool to assess the veracity of written statements. It is probably the most frequently used verbal assessment instrument worldwide, and is used as evidence in criminal courts in several countries in the world (Vrij, 2000). The core component of SVA is Criteria-Based Content Analysis (CBCA). CBCA comprises 19 different criteria (see the Appendix Appendix). CBCA-trained evaluators judge the absence or presence of each of these 19 criteria in a written transcript. The presence of each criterion strengthens the hypothesis that the account is based on genuine personal experience. In other words, truthful statements will have more of the elements measured by CBCA than false statements. A review of 37 CBCA studies showed strong support for the instrument. Out of 12 studies where total CBCA scores were calculated, 11 studies revealed significantly higher CBCA scores for truth tellers than for liars (Vrij, 2005). The only exception was found in Ruby and Brigham’s study (1998), and several explanations are possible for this deviant finding (Vrij, 2005). For example, Ruby and Brigham (1998) used raters who were trained for only 45 minute in CBCA scoring, and it is doubtful whether people could be considered “CBCA trained” after such a short training period. Despite the empirical support for CBCA, the system is not without its critics (Pezdek et al., 2004;
Rassin, 1999; Ruby & Brigham, 1997; Sporer, 1997; Wells & Loftus, 1991). For example, Pezdek et al. (2004) reported that CBCA scores are affected by event familiarity: Descriptions of familiar events are more likely to be judged true than are descriptions of unfamiliar events. Also, CBCA scores are positively correlated with age (Vrij, 2005).

Sporer (1997) focused his criticism on the atheoretical nature of CBCA. He pointed out that the psychological processes behind why differences in CBCA scores between genuine and fabricated accounts are likely to occur have not been specified. This is a crucial point as, without theoretical foundation, the boundaries of the instrument are unknown. That is, when are differences in CBCA scores likely to occur and under what circumstances are they unlikely to occur (Sporer, 1997). For example, Raskin and Esplin (1991) state that statements need to be of a certain length for CBCA to be applied but they do not make clear why. Also, CBCA was developed to evaluate statements from children who are witnesses or alleged victims in sexual abuse cases. Many authors still describe CBCA as a technique solely developed to evaluate statements made by children in sexual offence trials (Honts, 1994; Horowitz et al., 1997). Others, however, advocate expanding the use of the technique to evaluate the testimonies of adults who talk about issues other than sexual abuse (Köhnen, Schimossek, Aschemann, & Höfer, 1995; Porter & Yuille, 1996; Ruby & Brigham, 1997; Steller & Köhnken, 1989). Clearly, if the theoretical foundations of the instrument were established, then this could shed light on these issues.

Köhnen (1989, 1996, 1999) presented a theoretical foundation for CBCA and proposed that both cognitive and motivational factors influence CBCA scores. However, these theoretical assumptions have never been empirically tested so far. The present experiment provides such a test for the very first time.

With regard to cognitive factors, it is assumed that the presence of several criteria (criteria 1–13, see also the Appendix) are likely to indicate genuine experiences as they are typically too difficult to fabricate. Therefore, statements that are coherent and consistent (logical structure), whereby the information is not provided in a chronological time sequence (unstructured production) and that contain a significant amount of detail (quantity of detail) are more likely to be true. Regarding details, accounts are more likely to be truthful if they include contextual embeddings (references to time and space: “He approached me for the first time in the garden during the summer holidays”), descriptions of interactions (“The moment my mother came into the room, he stopped smiling”), reproduction of speech (speech in its original form: “And then he asked: Is that your coat?”), unexpected complications (elements incorporated in the statement that are somewhat unexpected, e.g. the child mentions that the perpetrator had difficulty with starting the engine of his car), unusual details (details that are uncommon but meaningful, e.g. a witness who describes that the man she met had a stutter), and superfluous details (descriptions that are not essential to the allegation, e.g. a witness who describes that the perpetrator was allergic to cats). Another criterion that might indicate truthfulness is when a witness speaks of details that are beyond the horizon of his or her comprehension, e.g. when he or she describes the adult’s sexual behaviour but attributes it to a sneeze or to pain (accurately reported details misunderstood). Finally, possible indicators of truthfulness are if the child reports details that are not part of the allegation but are related to it (related external associations, e.g. a witness who describes that the perpetrator talked about various women he had slept with and the differences between them), when the witness describes his or her feelings or thoughts experienced at the time of the incident (accounts of subjective mental state), or describes their interpretation of the perpetrator’s feelings, thoughts or motives.
during the incident (attribution of perpetrator’s mental state: “He was nervous, his hands were shaking”).

Other criteria (criteria 14–18) are more likely to occur in truthful statements for motivational reasons. Truthful persons will not be as concerned with impression management as deceivers. Compared to truth tellers, deceivers will be more keen to try to construct a report that they believe will make a credible impression on others, and will leave out information that, in their view, will damage their image of being a sincere person (Kohnken, 1999). As a result, a truthful statement is more likely to contain information that is inconsistent with the stereotypes of truthfulness. The CBCA list includes five of these so-called “contrary-to-truthfulness-stereotype” criteria (Ruby & Brigham, 1998): spontaneous corrections (corrections made without prompting from the interviewer (“He wore black trousers, no sorry, they were green”), admitting lack of memory (expressing concern that some parts of the statement might be incorrect: “I think”, “Maybe”, “I am not sure”, etc.), raising doubts about one’s own testimony (anticipated objections against the veracity of one’s own testimony: “I know this all sounds really odd”), self-deprecation (mentioning personally unfavourable, self-incriminating details: “Obviously it was stupid of me to leave my door wide open because my wallet was clearly visible on my desk”), and pardoning the perpetrator (making excuses for the perpetrator or failing to blame him or her, such as a girl who says she now feels sympathy for the defendant who possibly faces imprisonment).

The final criterion relates to details characteristic of the offence. This criterion is present if a description of events is typical for the type of crime under investigation (for example, a witness describes feelings that professionals know are typical for victims of, for example, incestuous relationships). It is difficult to see how this criterion fits into the cognitive and motivational theoretical framework. Raskin and Esplin (1991) pointed out that this criterion is not related to the statement itself but to the particular crime to which the statement refers to. They therefore argue that this criterion should not be included in the CBCA criteria list.

The cognitive and motivational theoretical framework fits nicely into other theoretical frameworks about deception. Back in 1981, Zuckerman, DePaulo, and Rosenthal argued that truth tellers and liars differ in behaviour and speech, amongst other factors, because of differences in cognitive load and motivation. In the present experiment, participants lied (n = 40) or told the truth (n = 40) about a certain event and were interviewed about that event Afterwards. Truth-telling participants had participated in a scenario involving money missing from a wallet, and truthfully reported what had happened during that event. Liars however were told about the event regarding money going missing from a wallet, but did not actually participate in it themselves, and instead were asked to lie and claim that they had been.

After this interview, the participants completed a questionnaire measuring their cognitive load and the extent to which they tried to control their speech. The interviews were audiotaped, transcribed and coded by CBCA raters. The following hypotheses were tested: (1) CBCA scores for truth tellers will be higher than CBCA scores for liars, (2) liars will experience more cognitive load than truth tellers, (3) liars will be more concerned with controlling their speech than truth tellers, (4) CBCA scores will be negatively correlated with cognitive load, and (5) CBCA scores will be negatively correlated with the desire to control speech. For reasons beyond the scope of this article, two different interview styles were employed, an “information-gathering technique” style and a “behaviour analysis interview” (BAI) style (Inbau, Reid, Buckley, & Jayne, 2001). Phase 1 of both techniques is identical: The participant is invited to recall in his or her own words in as much detail as
possible what happened during the event. The second phase of the information gathering
technique consists of several follow-up questions: (“You just mentioned that someone came
into the room. Can you describe that person in more detail?”). The second phase of BAI,
however, contains a series of questions, all described in Inbau et al. (2001), more about the
crime (taking money from a wallet) than about the event. For example: “Who would have
the best opportunity to have taken the money if they had wanted to?” (see the Method for
further details of the two interview techniques). Although Phase 2 of BAI will provide some
additional information about the event, it is likely to be less detailed than in the second
phase of the information-gathering technique. Therefore, a lower CBCA score in Phase 2
for BAI is likely to occur. Also, because less information is provided by participants in Phase
2 of BAI differences between liars and truth tellers in the second phase of BAI are less likely
to occur than in the second phase of the information gathering style. There are no reasons
to believe that the two styles will differ in terms of the cognitive load they will require and/or
in the desire of participants to control their speech.

Method

Participants

A total of 80 undergraduate students participated, of whom 44 were males and 36 were
female. Their age ranged from 18–46 years old and their average age was $M = 22.39$
(SD = 6.7).

Procedure

The experiment took place at a Students Union. Undergraduates were recruited under the
guise of participating in an experiment about “telling a convincing story” with the
possibility of earning £15. First, the participants signed an informed consent form. From
that moment events differed for the participants depending on whether they were allocated
to the truth-telling condition or deception condition. Random allocation was made to these
conditions. After entering the experimental room, participants allocated to the truth-telling
condition ($n = 40$) were approached by the experimenter who explained that they would earn
£5 for sure, and potentially an additional £10 if they were convincing, to participate in an
experiment investigating memory and the impression that people give of themselves. Once
the participants had consented to the study they were told that the experiment was run in
twos, and that the other participant was already waiting. They were taken up to a room in
the Student’s union where a confederate (Confederate 1, “Sam”) was alone sitting at a table
waiting, posing as another participant. Confederate 1 was played consistently by the same
person for all 40 truth tellers. (The following event is a slightly modified version of Vrij,
Akehurst, Bull, & Soukara, 2002.) Both “participants” (the participant and the confed-
erate) were then asked to play a game of Connect 4 and get to know each other and
someone would be along shortly. Confederate 1 would then start playing Connect 4 and
chatting to the participant, steering to one of a few predetermined subject areas. After a
minute, another confederate (Confederate 2) entered, apologized for interrupting and
claimed to need to wipe some mathematical equations off a whiteboard in the room. He/she
(Confederate 2 was changed from time to time) did so and then left the room after making
one of a selection of comments that had previously been scripted. After another minute
Confederate 1’s mobile phone rang and she excused herself and left the room (leaving the
participant alone) for a moment. She then returned and Connect 4 play recommenced.
After another minute Confederate 3 entered looking for a wallet (the person playing Confederate 3 was also varied from time to time). Confederate 3 also made comments as he/she looked for the wallet that were based on one of three scripts. The wallet had previously been placed in various places around the room. Confederate 1 (and normally the participant) would help Confederate 3 look for the wallet, which would then be found and Confederate 3 would express relief and leave. The experimenter then returned a minute later with Confederate 3 and explain that Confederate 3 claimed that a £10 note was missing from the wallet. Since Confederate 1 and the participant were the only two people to have been in the room with the wallet they would then be individually interviewed in the room next door by a police officer. Clearly, the participant was always chosen to be interviewed first. Afterwards the participant was taken off to another room to complete the questionnaire. After this they were all paid £15 regardless of how convincing they had been (but were informed that the police officer had considered their story to have been convincing).

Participants in the deception condition were recruited the same way. They were then taken upstairs to the same room that the truth tellers were taken to, set up in the same way, with the white-board containing the same mathematical equations. They were asked to take £10 out of the Wallet. They were then handed the following instructions to read:

“You entered the room to find Sam. The two of you are instructed by the experimenter to play Connect 4 for a while, which you do alone together. The other participant was sat where the experimenter was just sat and you sat where you are sitting now, and the wallet and Connect 4 game were on the table where you just saw them. You had a general conversation with the other participant as you played (e.g. about your courses/life as a student in Portsmouth/TV/the weather). Then someone else entered the room, made a comment about you playing the game, wiped those mathematical formulas off the board and then left. You continued playing the game for a few moments when Sam’s mobile phone rang and, clearly an important personal call, she excused herself and left the room leaving you alone for a minute or so. Then she returned and you both continued playing the game. Then someone else entered the room looking for his/her wallet. The person made several comments when s/he entered and had clearly been looking for it for a while and was very grateful to find it here in the room. You continue playing the game when the experimenter came back in, with the wallet-owner, and informs you and the other participant that some money had gone missing from the wallet and you are both to be interviewed. When asked about the events that have just recently occurred, explain in as much detail as you can, everything that has just happened in this room. Remember to give as much detail about everything in the room, everyone who has entered and/or exited the room, and everything that has been said and done in the room. The interviewer will then make a judgement, based on what you tell him, about whether or not you are telling the truth. If you manage to convince the interviewer that you did not take the money then you will receive an additional £10 to the £5 you are getting for participating in this study. However if you do not convince him that you are telling the truth then you will not receive the additional £10, and will have to give a written statement about what actually occurred. You now have a few minutes to familiarise yourself with this alibi before going into the next room to give your story to the interviewer. Remember to give as much detail as you can in order to make your story convincing.”

Any questions were then answered, and the participant was given a few minutes to read and absorb his/her story. When he/she was ready they were taken next door to be interviewed.
After the interview, as in the truth-telling condition, all participants completed a questionnaire, and were then told that the interviewer had considered their story to be convincing and so they were paid the £15.

Participants were interviewed by a uniformed British male police officer. The interviewer was blind to the condition (truth telling or lying) the participants were in. He started the interview by saying “£10 has gone missing from a wallet in the room next door and I have to find out whether or not it was you who took it”. He then asked several introduction questions including questions about the participant’s age, the course they study, and location they originally came from. These questions were asked for the purpose of giving the participants the opportunity to settle themselves down and were not designed for further analyses. Then the actual interview started with the question: “Could you tell me in as much detail as possible what happened when you were in the room next door with Sam just now. Mention all details, all conversations that took place, and give as much information as you can about everyone who entered the room, no matter how irrelevant it may seem”. The answer given by the participants is further referred to as “Phase 1”. In the information-gathering condition ($n=40$, 20 liars and 20 truth tellers) Phase 2 consisted of the following questions: (i) “Right, is there anything else that you can remember that you can tell me. Please think about any other details that you haven’t mentioned yet however irrelevant you think they might be”, (ii) “You just mentioned that someone came into the room who rubbed information off the board. Can you describe that person?”, (iii) “And also you mentioned that a person came in and claimed that the wallet was theirs. Can you describe them in detail?”, (iv) “And you said you played a game of Connect 4. Can you describe in detail how the game went?”, and (v) “Were there any conversations going on whilst you were in the room. Can you tell me everything you can recall being said?”. The answers to these five questions are further referred to as Phase 2. Phase 2 of the BAI ($n=40$, 20 liars and 20 truth tellers) comprised 15 questions (see Inbau et al., 2001, for further details), including “Do you know who did take the money?”, “Is there anyone other than yourself who you feel certain did not take the money?”, and “Tell me what would you stop from taking the money?”. The average number of words spoken by the participants in Phase 1 was $M=221$ (SD=115), and in Phase 2 $M=275$ words (SD =133). Veracity and Type of Interview had no effect on the number of words spoken in both Phases 1 and 2 (all $F$s <1.38 and all $p$s >0.24).

After the interview the police officer thanked them for their time and gave them a questionnaire. The participant completed the questionnaire and was then told by the experimenter (who pretended to check) that the police officer had considered that the participant’s story had been convincing and hence they were paid the full £15. This pretence was because it was decided to be more ethical that all participants should be paid the same amount, but if participants realized that they would be paid the full amount whether they were convincing or not, then the £10 bonus would no longer be an effective motivation to be convincing. The experimenter debriefed participants and answered any questions that they had. All interviews were audiotaped and transcribed.

*Questionnaire and CBCA scoring*

To measure participants’ tendency to control their speech the following five criteria were attempted: “During the interview I attempted to control what I was saying”, “During the interview I tried to say things that would sound honest”, “During the interview I put effort into trying to avoid saying certain things”, “During the interview I paid attention to what I
said”, and “During the interview I realized that others would analyse what I said in order to find out whether or not I was lying”. Answers to these criteria could be given on seven-point Likert scales ranging from (1) certainly not to (7) certainly. The answers were clustered into the “speech control” variable (Cronbach’s alpha = 0.71). To measure cognitive load, three more criteria were asked: “The interview required a lot of thinking (cognitive effort)”, “The interview was mentally difficult”, and “During the interview I had to concentrate a lot”. Again, answers could be given on seven-point Likert scales ranging from (1) certainly not to (7) certainly. The answers were clustered into the “cognitive load” variable (Cronbach’s alpha = 0.86). In order to check participants’ motivation (manipulation check) the following two criteria were proposed: “To what extent were you motivated to appear convincing during the interview?” (the answer could be given on a seven-point Likert-scale ranging from (1) definitely not to (7) definitely), and “During the interview I felt ‘duping delight’” (joy associated with meeting the challenge of convincing a police officer about the truthfulness of a story (Ekman, 1985)) (the answer could be given on a seven-point Likert scale ranging from (1) certainly not to (7) certainly). These two questions correlated significantly with each other, $r(80) = 0.39$, $p < 0.01$, and are clustered into the variable “motivation”.

The transcripts were scored by a CBCA expert and a second rater. The second rater received training in CBCA scoring by the British expert. First, the rater read several descriptions of the CBCA criteria provided by Raskin and Esplin (1991), Steller (1989) and Vrij (2000). Second, the expert explained each criterion under investigation (see below) and gave examples of each criterion. Third, both trainee and expert evaluated several example scripts together (from a different study), and then rated a few scripts individually. The trainee and expert compared their results and feedback was given by the expert rater. Fourth, the trainee received more transcripts and was asked to rate these at home. In a follow-up meeting, the results were evaluated and again, feedback was given by the expert. After this meeting both the rater and the expert felt that the trainee had been adequately trained and it was decided that the raters could commence their coding task. Both expert and second rater coded the scripts individually in their separate homes. The ratings took place on the basis of the written transcripts of the interviews and both raters were blind to the veracity condition (truth telling or lying) the participants were in. Criterion 10 “accurately reported details misunderstood” was not scored as it is specifically relates to young children talking about sexual crimes. Criterion 19 was not used either because, as discussed above, it is not concerned with the actual statement. The CBCA expert scored the frequency of occurrence of each criterion in each statement in Phases 1 and 2. Following common procedure, repeated information was not counted twice. Several questions in Phase 2 of the BAI did not reveal additional information about the event and therefore do not obtain CBCA scores, but several other questions did result in additional information that could subsequently be scored by the CBCA raters. After completing this frequency scoring, the expert then scored the presence of each criterion on five-point Likert scales, (1) = absent and (5) = strongly present, by using the frequency scores. A total CBCA score was calculated by adding the Likert scale scores of these 17 CBCA criteria. The total CBCA score could therefore range from 17 to 85, and this is the score used throughout the analyses. In order to check for inter-rater reliability, the second coder also conducted the frequency scoring ratings. Table I provides details of the inter-rater agreement (Pearson correlations) between the CBCA expert and the second rater for the 80 scripts in total and for the information-gathering and BAI interviews separately. Table I also provides a total CBCA score, which is the total frequency score of all 17 individual frequency scores
combined. Table I provides good and certainly acceptable inter-rater scores for all 17 variables. For the total sample, the two coders’ ratings correlated significantly at a $p < 0.01$ level for all 17 criteria. For the two types of interview separately most scores were significant at this level. The two CBCA total scores correlated very highly for each type of interview and for the two interview types combined.

### Results

#### Manipulation check

The vast majority of participants (84%) indicated that they were to some extent motivated to appear convincing during the interview (score > 4 on the seven-point Likert scale), but a minority of 39% felt to some degree a duping delight (score > 4 on the seven-point Likert scale). On the combined motivation scale, 62% obtained a score higher than the neutral value 4. In other words, most participants appeared to be motivated to appear convincing in the experiment. This motivation (on the combined scale) was not affected by the Veracity condition they were in, $F(1,76) = 2.29$, NS, or the Type of Interview they were exposed to, $F(1,76) = 0.51$, NS. Neither was the Veracity × Type of Interview interaction effect significant, $F(1,76) = 0.39$, NS. When we examine the two individual items separately, the answer to the motivated to appear convincing variable was not affected by Veracity, Type of Interview or the interaction between these two factors (all $F$s < 1.60, all $p$s > 0.21). Regarding duping delight one significant effect emerged. Truth tellers felt more duping delight ($M = 4.40$, SD = 1.6) than liars ($M = 3.63$, SD = 1.7), $F(1,76) = 4.04$, $p < 0.05$, $\eta^2 = 0.05$, which is interesting as Ekman (1985) who was the first to associate duping delight with deception predicted that truth tellers would be less rather than more delighted than liars. His prediction is not supported in the present experiment.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Total ($n=80$)</th>
<th>Info-gathering ($n=40$)</th>
<th>BAI ($n=40$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.35**</td>
<td>0.39*</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.47**</td>
<td>0.45**</td>
<td>0.50**</td>
</tr>
<tr>
<td>3</td>
<td>0.95**</td>
<td>0.94**</td>
<td>0.88**</td>
</tr>
<tr>
<td>4</td>
<td>0.90**</td>
<td>0.88**</td>
<td>0.94**</td>
</tr>
<tr>
<td>5</td>
<td>0.42**</td>
<td>0.31</td>
<td>0.57**</td>
</tr>
<tr>
<td>6</td>
<td>0.97**</td>
<td>0.98**</td>
<td>0.83**</td>
</tr>
<tr>
<td>7</td>
<td>0.47**</td>
<td>0.49**</td>
<td>0.50**</td>
</tr>
<tr>
<td>8</td>
<td>0.64**</td>
<td>0.68**</td>
<td>0.39*</td>
</tr>
<tr>
<td>9</td>
<td>0.60**</td>
<td>0.67**</td>
<td>0.31</td>
</tr>
<tr>
<td>11</td>
<td>0.38**</td>
<td>0.29</td>
<td>0.62**</td>
</tr>
<tr>
<td>12</td>
<td>0.77**</td>
<td>0.74**</td>
<td>0.85**</td>
</tr>
<tr>
<td>13</td>
<td>0.84**</td>
<td>0.88**</td>
<td>0.78**</td>
</tr>
<tr>
<td>14</td>
<td>0.64**</td>
<td>0.66**</td>
<td>0.60**</td>
</tr>
<tr>
<td>15</td>
<td>0.78**</td>
<td>0.75**</td>
<td>0.77**</td>
</tr>
<tr>
<td>16</td>
<td>0.44**</td>
<td>0.48**</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.61**</td>
<td>–</td>
<td>0.55**</td>
</tr>
<tr>
<td>18</td>
<td>0.49**</td>
<td>–</td>
<td>0.32*</td>
</tr>
<tr>
<td>Total CBCA score</td>
<td>0.96**</td>
<td>0.95**</td>
<td>0.93**</td>
</tr>
</tbody>
</table>

* $p < 0.05$, ** $p < 0.01$. 

A. Vrij and S. Mann
Hypotheses testing

In order to test Hypothesis 1 (truth tellers obtain higher CBCA scores than liars) a 2 (Veracity) × 2 (Type of Interview) MANOVA was conducted with the total CBCA scores for Phases 1 and 2 as dependent variables. At a multivariate level, the Veracity main effect, $F(2,75) = 3.48, p < 0.05, \eta^2 = 0.09$, the Type of Interview, $F(2,75) = 29.14, p < 0.01, \eta^2 = 0.44$, and the Veracity × Type of Interview interaction effect, $F(2,75) = 3.27, p < 0.05, \eta^2 = 0.08$, were all significant. Univariate tests regarding the Veracity factor showed that in both Phase 1 (truth tellers: $M = 29.95, SD = 5.6$ vs liars: $M = 27.88, SD = 2.6$), and in Phase 2 (truth tellers: $M = 27.22, SD = 5.5$ vs liars: $M = 25.20, SD = 2.7$) truth tellers obtained higher CBCA scores than liars (Phase 1: $F(1,76) = 4.30, p < 0.05, \eta^2 = 0.05$, Phase 2: $F(1,76) = 6.66, p < 0.05, \eta^2 = 0.08$), which supports Hypothesis 1. Univariate tests for Type of Interview revealed no effect for Phase 1, $F(1,76) = 0.00, NS$, but a significant effect for Phase 2, $F(1,76) = 37.06, p < 0.01, \eta^2 = 0.33$. Unsurprisingly, participants obtained a higher CBCA score in the information-gathering interview ($M = 28.60, SD = 4.8$) than in the BAI ($M = 23.83, SD = 2.2$). Univariate tests for the Veracity × Type of Interview interaction did not reveal a significant effect for Phase 1, $F(1,76) = 0.03, NS$, however, the effect for Phase 2 was significant, $F(1,76) = 4.56, p < 0.05, \eta^2 = 0.06$. In the information gathering condition truth tellers obtained a higher CBCA score in Phase 2 ($M = 30.45, SD = 5.7$) than liars did ($M = 26.75, SD = 2.6$), $F(1,38) = 6.95, p < 0.05, \eta^2 = 0.16$. In Phase 2 of the BAI interview, liars and truth tellers obtained similar CBCA scores, $F(1,38) = 0.25, NS$.

In order to test Hypotheses 2 (liars will experience more cognitive load than truth tellers) and 3 (liars will be more concerned with controlling their speech than truth tellers) a 2 (Veracity) × 2 (Type of Interview) MANOVA was conducted with speech control and cognitive load as dependent variables. At a multivariate level, the analysis revealed a significant effect for Veracity, $F(2,75) = 11.40, p < 0.01, \eta^2 = 0.23$. The main effect for Type of Interview, $F(2,75) = 2.90, NS$, and the Veracity × Type of Interview interaction effect, $F(2,75) = 1.46, NS$, were not significant. Univariate tests regarding the Veracity factor revealed significant results for both speech control, $F(1,76) = 17.61, p < 0.01, \eta^2 = 0.19$, and cognitive load, $F(1,76) = 11.32, p < 0.01, \eta^2 = 0.13$. As was predicted in Hypothesis 2, liars experienced more cognitive load in their interviews ($M = 4.42, SD = 1.1$) than did truth tellers ($M = 3.50, SD = 1.4$). Hypothesis 2 is therefore supported. In alignment with our predictions, liars also tried to control their speech more ($M = 5.24, SD = 0.9$) than truth tellers ($M = 4.30, SD = 1.1$), which supports Hypothesis 3.

In order to test Hypotheses 4 (CBCA scores will be negatively correlated with cognitive load) and 5 (CBCA scores will be negatively correlated with the desire to control speech) Pearson correlations were conducted. The CBCA score was correlated with the speech control and cognitive load scores. The results are provided in Table II (top). As can be seen in Table II, speech control was not significantly correlated with CBCA scores. Hypothesis 5 is therefore rejected. Cognitive load was correlated with CBCA scores, but only for truth tellers and only in Phase 2 (the correlation for Phase 1, $r(80) = 0.29, p = 0.07$, was marginally significant). However, the correlation revealed that cognitive load was positively correlated with CBCA scores whereas a negative correlation was predicted. Therefore, Hypothesis 4 is also rejected. As discussed above cognitive load and speech control were predicted to be correlated with specific CBCA variables, rather than with the total CBCA score. We therefore divided the total CBCA scores into three sub-elements: (i) the general criteria 1–3, (ii) the cognitive criteria 4–13, and (iii) the motivation criteria 14–18.
Perhaps cognitive load would correlate as predicted with the general and cognitive criteria and speech control with the motivation criteria. The correlations for these three clusters are also presented in Table II (middle). Even at this, more detailed level, no significant correlations emerged for speech control. For cognitive load, several significant correlations emerged but they were similar to the general pattern and contradict what we expected: Cognitive load was positively correlated with the general characteristics in Phases 1 and 2 and with the cognitive variables in Phase 2. Hypotheses 4 and 5 therefore need to be rejected.

Finally, we explored whether CBCA scores would be correlated with participants’ motivation to do well in the task (the manipulation check). One could argue that motivated participants would put more effort into giving a complete recall, which subsequently will result in a higher CBCA score. The correlations between total CBCA scores and motivation (the “motivation” variable) are presented in Table II (bottom). Indeed, CBCA scores were positively correlated with motivation in Phase 1.

**Discussion**

We tested the underlying assumptions of CBCA that fabricators experience more cognitive load and are more concerned with what they say than truth tellers. We found support for the
hypotheses that liars experience more cognitive load and try harder to control their speech. We also found that CBCA scores for liars were lower than those for truth tellers. However, CBCA scores, experienced cognitive load and speech control were not correlated with each other in the predicted way. Speech control was not correlated with CBCA scores at all. In other words, whatever the participants (i) tried to say or (ii) tried to avoid saying in order to appear honest, it did not affect their CBCA scores. It could be that the speech characteristics that CBCA experts believe liars would avoid including in their reports (criteria 14–18) are not obvious to most liars and hence they don’t think about deliberately avoiding to include these characteristics when they try to appear honest. However, if that is the case then it is difficult to argue that, when these CBCA criteria are absent in deceptive statements, it is the result of liars’ speech control.

Cognitive load was correlated with CBCA scores but in an unpredicted way. Cognitive load and CBCA scores were not correlated for liars, but for truth tellers a positive correlation emerged (the more cognitive load the participants experienced, the higher their CBCA scores were). CBCA theory predicted a negative correlation between cognitive load and CBCA scores. A positive correlation could be rather easily explained. Recalling a witnessed event in a lot of detail is cognitively difficult, so it is no surprise that truth tellers, when they succeed in having done this, report that they found the task difficult. However, a rich account also leads to high CBCA scores. Whatever the explanation for the positive cognitive load–CBCA scores correlation is, it means that increased cognitive load cannot be used to explain why liars obtained lower CBCA scores than truth tellers in the present study.

CBCA scores were positively correlated with the extent to which participants were keen to perform well in the present study. This implies that those who are motivated to perform well succeeded in doing this and obtained high CBCA scores. However, this still doesn’t tell us why motivation helped or what highly motivated participants did in order to obtain high CBCA scores.

In summary, we failed to uncover the underlying theoretical principles of CBCA scores. This doesn’t mean that CBCA cannot be used as a practical tool to distinguish between truth and fabrication. However, a lack of theoretical foundation is unfortunate as this makes it unclear what the possibilities and restrictions of the instrument are. That is, accurate predictions when an instrument can be used and when it cannot be used can only be made on the basis of sound theoretical principles. Further research into these theoretical principles is therefore highly desirable. Such research could also focus on young children, as the CBCA instrument is developed particularly for this age group. Our present findings with adults do not rule out that the cognitive load and speech control hypotheses could be supported in a younger age group.

Acknowledgements

This study was sponsored by a grant from the Economic and Social Research Council (RES-000-23-0292).

Note

1 In order to avoid all truth tellers giving the same descriptions, different confederates were used, wearing different clothing. For the same reason, the remarks made by the confederates and the topics discussed during the game of Connect 4 varied.
References


Appendix. Content criteria for statement analysis

*General characteristics*
1. Logical structure
2. Unstructured production
3. Quantity of details
Specific contents
4. Contextual embedding
5. Descriptions of interactions
6. Reproduction of conversation
7. Unexpected complications during the incident
8. Unusual details
9. Superfluous details
10. Accurately reported details misunderstood
11. Related external associations
12. Accounts of subjective mental state
13. Attribution of perpetrator’s mental state

Motivation-related contents
14. Spontaneous corrections
15. Admitting lack of memory
16. Raising doubts about one’s own testimony
17. Self-deprecation
18. Pardoning the perpetrator

Offence-specific elements
19. Details characteristic of the offence

Adapted from Steller and Köhnken (1989).