Applying criteria based content analysis to assessing the veracity of rape statements

Nathan Critchlow
Applying criteria based content analysis to assessing the veracity of real rape statements

**ABSTRACT** Rape is a serious crime and one of the most difficult to investigate due to a lack of witnesses. Therefore understanding veracity of a rape statement is of paramount importance. Police forces currently have no procedure to assess statement veracity therefore the aim of this investigation was to assess the application of Criteria Based Content Analysis (CBCA) to adult rape statements.

This investigation applied (CBCA) to 28 rape statements equally divided into categories of veracity (genuine, credible, non-credible and confirmed false allegations). According to the Undeutsch (1982) hypothesis, a genuine and credible statement will differ in content to a false or non-credible statement, suggesting CBCA could potentially be an investigative police tool.

ANOVA's and post-hoc t-tests demonstrating genuine (M=18.86) and credible (M=15.57) statements showed significantly higher criteria presence than non-credible (M=6) and false statements (M=6.29), but no significant difference between genuine and credible statements, or between non-credible and false statements. Furthermore the percentage of criteria present in genuine and credible statements indicates CBCA can be used to tackle adult rape narrative. Results indicate CBCA could be used to assess veracity of adult rape statements, and have application as a police investigative tool.

<table>
<thead>
<tr>
<th>KEY WORDS:</th>
<th>CBCA</th>
<th>FALSE ALLEGATION</th>
<th>STATEMENT VALIDATION</th>
<th>RAPE</th>
<th>NFA</th>
</tr>
</thead>
</table>
Background

Rape is regarded as one of the most serious and distressing crimes a person can be subjected to and furthermore, due to the inherent lack of witnesses, it is also one of the most difficult for the police to investigate (Gibson, 1998; Parker & Brown, 2000). Research within this domain has focused on a wide array of topics including: criticism of police scepticism of certain accounts of rape (Adler, 1987; Gregory & Lees, 1996; Smith, 1989; Temkin, 1999); police policies (Jones, Newburn & Smith, 1994); police procedures (Temkin, 1987); police attitudes regarding victims (Brown & King, 1998) and victim attitudes towards the police (Lees, 1997). However, it appears there is little research focusing on assessing the veracity of a rape allegation - in particular through analysis of the statements and whether it is possible to use scales as an investigative tool for assessing whether a statement is a false or genuine allegation.

False allegations pose significant problems for police officers investigating sexual assault and rape. Moreover, any investigating officer who wrongly accuses a victim of fabricating an allegation may cause additional problems, compound the victim’s psychological and emotional trauma and seriously hamper the relationship between the victim and the investigating officers (McDowell & Hibler, 1987; Marshall & Alison, 2006). Conversely, the failure of police to accurately detect false allegations may be viewed as a waste of police time or, more significantly, lead to the wrongful conviction of an innocent individual (MacDonald, 1971; Marshall & Alison, 2006).

One of the most challenging elements of investigating false allegations of rape is the contrasting reports of how frequent false allegations are in society. Feminist theorists argue a false allegation doesn’t exist, and allegations which are regarded as false are, simply, those not believed by the investigating police officers (Russell, 1984). Empirical evidence has proven equally inconclusive with O’Reilly (1984) arguing the frequency of false allegations of rape is as low as 0.25%, Theilade & Thomsen (1986) claiming around 10% are false and remarkably Kanin (1994) claims it could be as high as 41%. These contrasting statistics are, to an extent, explained through the lack of a standardized methodology to conceptualize the problem that, in turn, affects between studies verification, quality of repeated methodologies and different law enforcement terminology used (Parker & Brown, 2000). It is also important to consider cases in which ‘no further action’ (NFA) is taken. This is a broad category and can arise for a number of reasons including: victims resistant to attending court; the Crown Prosecution Service refusing to prosecute, lack of corroborating evidence or contradictory medical evidence (Marshall & Alison, 2006). Nonetheless, it is critical to acknowledge, just because a case is classified as NFA, it doesn’t mean the allegation is false; it could suggest something has occurred, prompting the decision to report the matter to the police in the first instance, or it may be the allegation is genuine, but discontinued by the victim to protect psychological well-being.

Additionally, these methodological limitations are also compounded by the range of different theories as to why false allegations arise or, as some theorists argue, whether they actually exist at all (Adshead, 1996; Williamson, 1996). Research into this field is vast and varied. A cursory review gives the reader an
idea of the breadth of studies available, encompassing a wide range of social factors such as; hiding infidelity, social pressures, celebrities, unpaid prostitution (Aiken, Burgess & Hazelwood, 1995), alibi for feared pregnancy, revenge, attention and sympathy seeking (Kanin, 1994) or financial motivation (West, 1987). Parker & Brown (2000) argue for two major theoretical types of false allegation. The first of these is described as a deliberate and deceptive act in which the false allegation is intentionally fabricated. The second theory posits false allegations are false memories arising either as a result of delusional states, flawed memories or suggestive influences (Parker & Brown, 2000). A third theory of false allegation of rape purports a false allegation may arise due to coercion (Gibson, 1998).

Police forces currently have no standard procedure for assessing the veracity of rape statements and, consequently, investigating officers often have to make subjective judgements based upon ‘shorthand cues’ (Marshall & Alison, 2006), heuristics and a reliance on wisdom or experience (McDowell & Hibler, 1987; Gibson, 1998). However, one of the problems associated with these methods is, as some research has indicated, police officers often hold the same beliefs as lay persons in relation to deceptive indicators e.g. body language (Akehurst, Köhnken, Vrij & Bull, 1996). This is, in part, supported by several other studies which have found police officers are no better at detecting deception than lay persons (Ekman & O'Sullivan, 1991; Ekman, O'Sullivan, & Frank, 1999) and detection rates are typically similar to chance (Ekman & O'Sullivan, 1991). A number of researchers have attempted to provide objective methods of detecting deception such as forced choice testing (Miller, 1999), body language (Vrij, 2000) and non-verbal behaviour (Scherer and Ekman, 1982). However, these are all subject to the above criticism of investigating officers erroneous beliefs, or are difficult to apply in a rape scenario as the victim’s focus, concentration and level of distress would make it difficult to obtain a clear cut result (McDowell, 1992).

However, a similar area of inquiry which may offer insight into understanding veracity and deceptive intentions in statements, would be the use of statement validation assessment (SVA) and, in particular, criteria based content analysis (CBCA) presently being used in assessing the veracity of statements relating to child abuse cases. The use of statement validation and CBCA was developed based on the Undeutsch (1982; 1984) hypothesis. This hypothesis states a truthful statement will differ in content, quality and characteristics from a false allegation based on fantasy (Undeutsch, 1982) and claims it’s possible to make assumptions about an individual’s psychological state by systematically identifying specified characteristics in a statement (Holsti, 1969).

A comprehensive review of field studies, assessing the forensic credibility of criteria based content analysis (CBCA), was conducted by Vrij (2005). There is considerable research looking at CBCA in artificial settings but, to outline the scale’s credibility for police investigations, it’s important to look at its applications in field research. The first example of CBCA being applied in a field study was conducted by Esplin et al. (1988). The results in this study were startling with confirmed cases obtaining a mean score of 24.8 (maximum 38) and doubtful cases scoring a mean of 3, confirming the Undeutsch (1982) hypothesis. However, this study was criticised as only one evaluator scored the transcripts.
Moreover, decisions on the veracity of the statements analysed was not established through case facts (Vrij, 2005) and the age ranges of the sample were large (Wells & Loftus, 1991). These criticisms were tackled by Boychuk (1991) who had looked at child abuse statements which had been rated independently by three researchers, and with a narrower sample which was divided into confirmed cases, doubtful and likely cases, established using case facts. The results of Boychuk (1991) were equally encouraging and, although not obtaining results quite as significant as Esplin et al. (1988), all the differences identified were still found to be in the expected direction i.e. confirmed cases scored higher CBCA means than doubtful cases. These findings were further supported in an important field study conducted by Lamb, Sternberg & Esplin et al. (1997). They analysed ninety-eight alleged child abuse statements in which there was physical evidence of contact between the child and the accused person and, for which, there was corroborating evidence. Once again the results were found in the expected direction, with confirmed cases scoring higher CBCA scores than false allegations.

These findings show compelling evidence for the use of CBCA, particularly when compared with the limited success of other means of detecting deception, in which the result is often no better than chance (O’Sullivan & Ekman, 1991). Vrij’s (2005) meta-analysis shows there is clear evidence to support the Undeutsch hypothesis (1982; 1984) and the application of CBCA as a method of discriminating between real, experienced events and those which arise as a result of fantasy, coercion or suggestion (Lucas & McKenzie, 1999).

Although certain researchers have advocated CBCA as a technique developed solely to test veracity in child abuse statements (Honts, 1994; Horowitz, Lamb and Esplin et al., 1997), there is an additional domain of research which argues the CBCA technique has application in the testimonies of adults (Kohnken, 1995; Porter & Yuille, 1996; Ruby & Brigham, 1997; Steller & Kohnken, 1989). One example of a field investigation which has utilized CBCA and statement validation analysis (SVA) to assess the veracity of adult rape statements is Parker & Brown (2000). They analysed forty-three interview statements of rape cases. Statements were collected using the cognitive interview (CI) and analysed using the CBCA methodology. As with studies of child abuse statements, as outlined in Vrij (2005), Parker & Brown (2000) also found when the CBCA model was applied it became possible to differentiate between genuine and false allegations of rape, with confirmed allegations scoring a higher CBCA score than false allegations, and claimed CBCA was able to indicate whether ‘unsubstantiated cases’ were genuine or false. Parker & Brown (2000) concluded the CBCA scale could provide added investigative value to police investigations. However, this study received criticism because only three categories of veracity were used to label the statements i.e. credible (true), indeterminable (unsubstantiated) and non-credible (false). This was an understandable criticism as the study failed to take into account the possibility some unsubstantiated cases were likely to be genuine whereas others were likely to be false. A second criticism levied at this study was that the categories established lacked ground-truth. For example, the criterion for establishing what constitutes a statement was ‘convincing evidence of rape’ and ‘corroboration in the legal sense and with either a suspect being identified or charged’ (Parker & Brown, 2000). Vrij (2005) argues these criteria are far too vague and are not
based on case facts. Furthermore, the majority of cases were only marked by one individual and it's unclear whether they were familiarized with case facts or whether s/he had any background knowledge of the case.

**Aims**

Taking account of the previous research outlined, this investigation had two primary aims. The first aim of this study was to investigate further the initial findings of Parker & Brown (2000), by dividing the unsubstantiated category in credible and non-credible and to base the genuine and false categories on ground truth as contained in the case facts. The second aim for this investigation was to establish to what extent CBCA could be used be assess the complexity of adult narrative in rape cases and whether the CBCA scale could be applied as a standardized procedure in police practice as a veracity verifying procedure when investigating rape.

**Hypotheses**

1: CBCA scores for genuine statements will be significantly higher than false statements.

2: CBCA scores for credible statements will be significantly higher than non-credible statements.

3: CBCA scores for credible statements will not be significantly different than genuine statements.

4: CBCA scores for non-credible statements will not be significantly different than false statements.

**Method**

**Design**

The design for this investigation deployed an independent groups design. This investigation operationalised one independent variable (IV). The IV was the defined veracity of the statement. The IV had four levels which referred to the veracity of the statement: genuine statement; credible statements; non-credible statements and false statements.

The dependent variable (DV) was the score the statement obtained when CBCA was applied.

**Sample**

The sample consisted of twenty-eight rape statements (N=28). This sample was recruited via stratified sampling and was extracted from a larger sample database (N=125) from Nixon (in press). Furthermore, permission was obtained from Merseyside Police Forces Rape and Sexual Assault (RASA) investigation team for the investigator to access the rape statements. All the statements were collected by trained members of the Merseyside Police Force using the principle of the CI (Geiselman & Fisher, 1989), and no contact was made between the
researcher and any of the rape victims or offenders. Moreover, in cases in which multiple statements were collected it was only the first statement which was analysed using the CBCA. The sample (N=28) was further sub-divided equally into four different conditions, as defined by the veracity of the statements. The four conditions were as follows: genuine statements; confirmed false statements; a credible statement and a non-credible statement. There has been much debate in finding an ecologically valid method defining statement veracity (McLean, 1979; McDowell, 1992), and it was considered critical to establish sound categories. Consequently, the categories for this investigation are defined as follows:

**Genuine (N=7)**

A case in which a ground truth or case fact has been established e.g. medical examination or CCTV evidence was available, and the case had been forwarded to the Crown Prosecution Service (CPS) and the outcome was a conviction in court.

**False (N=7)**

A case in which a ground truth had been established and the allegation has been proven false by the presence of additional forensic evidence e.g. C.C.T.V, or conflicting medical evidence was available or the victim who had made the allegation had given written confirmation to the police that their allegation was false.

**Credible (N=7)**

A case which had passed the ‘threshold test’ for evidence within police procedure and forwarded to be considered for charge by the Crown Prosecution Service (CPS) but which then decided to take no further action (NFA) or a case which was NFA’d by the victim. These cases had, additionally, been independently rated by two Detective Inspectors (DI’s) of Merseyside Police and an independent researcher, and for which the majority verdict and the statement were genuine.

**Non-credible (N=7)**

A case in which a Merseyside Police Force evidence review officer (ERO) had established there was insufficient evidence to support the allegation or the victim had asked for NFA to be taken. Furthermore, these cases had also been independently rated as being false or suspicious by two Merseyside Police Force Detective Inspectors (DI’s) and an independent researcher.
Apparatus

There are several different versions of the CBCA posited in psychological literature. This investigation used the nineteen-point scale (see Appendix 1) developed by Steller & Kohnken (1989) and cited in Vrij’s (2005) meta-analysis of CBCA literature. This scale contains nineteen different criteria which are then divided into three categories. These are as follows (see appendix 2 for definitions):

General Characteristics (Items 1-3)

Focus on non-specific elements of the statement e.g. ‘logical structure’ and ‘quantity of detail’. The presence of these are said to indicate truthfulness as they are difficult to fabricate (Vrij, 2005).

Specific Criteria (Items 4-13)

Concentrates on the specific details of the statements. Examples include ‘contextual embedding’ and ‘superfluous details’. The presence of these criteria is said to indicate truthfulness as they are also difficult to fabricate (Vrij, 2005).

Motivation Released Content (Items 14-18)

Concentrates on assessing the motivation of the individual writing the statement e.g. ‘admitting lack of memory’. The presence of these is said to indicate truthfulness as truth-tellers are less concerned with impression management when compared to deceivers (Vrij, 2005).

Offense Specific Elements (Item 19)

This item evaluates to what extent the events described are typical of the crime under investigation (Vrij, 2005).

Coding

There is no standardized methodology for coding the presence of items on the CBCA scales (see Vrij, 2005 for meta-analysis of CBCA application). For the purpose of this investigation the scoring criteria used by Parker & Brown (2000), Esplin et al. (1988) and advocated by Steller (1989) was used. This is a dichotomous three point scoring system. An item scored 0 if it was absent in the statement, it scored 1 if the item was present in the statement and it scored 2 if the item was strongly evident in the statement. For example, when a statement contained no detail it was scored 0, if there was a moderate amount it was given a score of 1 and, if there was a large amount of detail it was given a score of 2. Jones (1994) argues a dichotomous scoring system is beneficial as it creates a continuum of certainty.

Therefore, in this investigation, the minimum score a statement could have obtained was 0 (indicating a false statement) and the maximum which could be obtained was thirty-eight (indicating truthfulness).
Inter-Rater Reliability

To avoid investigator bias in the investigator's scoring of CBCA criteria a random sample (N=8) was independently scored by the investigator's supervisor, currently employed as both a chartered forensic psychologist for Merseyside Police Force and a senior lecturer at Manchester Metropolitan University. These were correlated to ensure reliability. Any major discrepancies were discussed by the investigator and the independent rater until a final decision was agreed.

Procedure

Permission was requested and granted by Merseyside Police Force and Nixon (in press) for the investigator to access a sample (N=125) of rape statements. All the statements were collected by trained sexual assault officers of Merseyside Police Force and all the interviews were structured and based on the principles of the CI (Geiselman & Fisher, 1989) and followed police national guidelines (CPTU, 1992). The sample was divided into four separate conditions. The investigator was situated in a secure office in Mather Avenue Police Station of Merseyside Police Force. Each statement was given to the investigator individually and, having been read twice, they were subsequently rated using the CBCA scoring criteria (see Appendix 1) and the CBCA scoring sheet (see Appendix 3). Once the investigator had finished scoring the statements, a random sample (N=8) was given to the independent investigator for scoring. Once this was complete, the scores were correlated to assess the internal validity of the scoring. Any major discrepancies were discussed between the investigator and the independent rater until agreed upon.

Ethics

Manchester Metropolitan University (MMU) approved ethics forms AEAF (see Appendix 4) and ECF (see Appendix, 5) were completed and signed by the investigator's supervisor and MMU examination board. By completing this ethics form the investigator agreed to abide by the British Psychological Societies (BPS) and universities ethical guidelines for research. The investigator and supervisor were also vetted by Merseyside Police Force and signed the Data Protection Act (1998) and Official Secrets Act (1989).

Analysis of the statements took place solely at Mather Avenue Police station of Merseyside Police Force. The police force representatives, investigator and the supervisor agreed none of the statements were to be removed by the investigator from police premises in either hard or electronic form, nor was any of the sensitive information to be sent via external e-mail. All data taken off Merseyside Police premises was converted into numerical form to maximise security. The nature of the statements was incredibly sensitive and it was of paramount importance this information was not made available in the public domain. Due to the confidential nature of the data no names or details of the rape cases, the victims or offenders analysed were revealed, at any point, in this report and the crime numbers are only used to identify cases in the raw data.
The researcher was fully briefed by both the supervisor and Merseyside Police’s RASA team of the sensitive and distressing nature prior to engaging with the rape statements. Moreover, the investigator remained under the supervision of trained members of the RASA team whilst analysing the statements. This was arranged to limit and tackle any risk, distress, psychological harm or questions the researcher may have had in relation to the rape statements. Following the completion the investigator was fully debriefed.

Neither the investigator nor the research supervisor was involved, in any way, with collection of the initial statements. These were conducted by fully qualified members of Merseyside Police Force and, as explained previously, the interviews were based on the CI principle (Geiselman & Fisher, 1989). Participants, under the age of eighteen were not interviewed without the presence of a parent or legal guardian. If appropriate, permission was obtained by Merseyside Police to access victims’ medical records. Victims of the rapes were offered medical attention and counselling services, to minimise the emotional and psychological trauma.

**Analysis**

To investigate whether there was a difference between the scores for the different veracity of the statements, the total means and standard deviations were calculated.

A one-way independent ANOVA was conducted to look for a main effect of statements’ veracity.

Six independent post-hoc t-tests were conducted which helped to ascertain which categories were significantly different from each other.

All veracities were analysed for the percentage presence of CBCA criteria in order to find out to what extent a particular criteria was present and whether CBCA could grasp adult rape narrative.

Inter-rater reliability was calculated using Pearson’s correlation.

**Results**

**Descriptive Statistics**

**Table1**

Means and standard deviations for the CBCA scores obtained for each statement veracity.

<table>
<thead>
<tr>
<th></th>
<th>Genuine</th>
<th>Credible</th>
<th>Non-Credible</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Genuine</td>
<td>18.86</td>
<td>3.08</td>
<td>16.57</td>
<td>3.60</td>
</tr>
<tr>
<td>Credible</td>
<td>16.57</td>
<td>3.60</td>
<td>6.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Non-Credible</td>
<td>6.00</td>
<td>2.00</td>
<td>6.29</td>
<td>1.38</td>
</tr>
<tr>
<td>False</td>
<td>6.29</td>
<td>1.38</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To examine whether the CBCA scale (Steller & Kohnken, 1989) could indicate veracity of adult rape statements, statements (N=28) were divided into four categories. As depicted by table one, the descriptive statistics show the genuine statements group (N=7) scored the highest on CBCA criteria (M=18.86, SD=3.08), whereas false statements (N=7) scored low (M=6.29, SD=1.38). Furthermore the credible statements (N=7) also scored highly (M=16.57, SD=3.60) and non-credible statements (N=7) scored the lowest (M=6.00, SD=2.00). This can be seen in figure one (below).

![Figure 1: Mean CBCA scores for the Genuine, Credible, Non-Credible and False Statements](image)

**One-way Independent ANOVA**

Scores for each criterion were subjected to a one-way independent ANOVA to assess whether the difference between scores was significant.

**Veracity of statement**

A main effect for the veracity of the statements was found to be significant; F (3, 27) = 44.98, p<0.001. This indicates a significant difference on CBCA scoring for the four different veracities of rape statements. Overall the means indicate the scores for genuine (M=18.86) and credible (M=16.57) were higher than false (M=6.29) and non-credible (M=6).

**Post-Hoc T-test**

Six independent t-tests were conducted to ascertain where the significant difference lay between the groups, with CBCA scoring as the DV. Levene’s tests of Equality of Variances were insignificant in all t-tests, therefore equal variances were assumed.
Genuine and False

An independent t-test indicated the mean CBCA score for genuine statements (M= 18.86, SD= 3.08) were significantly higher, t(12)=9.86, p<0.001, than CBCA scores for false statements (M=6.29, SD= 1.38).

Genuine and Non-Credible

An independent t-test further showed CBCA scores for genuine statements (M=18.86, SD=3.08) were significantly higher, t(12)=9.27, p<0.001, than CBCA scores for non-credible statements (M=6, SD=2).

Genuine and Credible

An independent t-test demonstrated there was no significant difference, t(12) =1.28, p>.05, between the CBCA scores for genuine statements (M=18.86, SD=3.08) and credible statements (M=16.57, SD= 3.60).

Credible and False

An independent t-test demonstrated CBCA scores for credible statements (16.57, SD=3.59) were significantly higher, t(12)=7.06, p<0.001, than CBCA scores for false statements (M=6.29, SD=1.38).

False and Non-credible

An independent t-test demonstrated there was no significant difference, t(12) =.31, p>.05, between the CBCA scores for false statements (M=6.29, SD=1.38) and non-credible statements (M=6, SD=2).

Non-Credible and Credible

An independent t-test demonstrated the mean CBCA score for credible statements (M=16.57, SD=3.60) was significantly higher, t(12)=6.79, p<0.001, than non-credible statements (M=6, SD=2).

Criteria Presence

To assess whether CBCA can accurately grasp adult rape narrative, Table 2 assesses the percentage presence of the CBCA criteria in genuine and credible rape statements.

Table 2 and figure 2 illustrate 16 of the 19 criteria were present in at least one genuine or credible statement, with 13 of these 16 criterion present in more than half of genuine and credible statements. The criteria most present in rape statements are; quantity of details (100%); contextual embedding (100%), description of interactions (100%), accounts of subjective mental state (100%) and details characteristic of crime (100%). The criteria which did not appear in any genuine or credible statements were; pardoning perpetrator (0%), accurately reported details misunderstood (0%), raising doubts about one’s testimony (0%). The high percentage of criteria detected in genuine and credible
statements indicates the CBCA scale is capable of analysing adult rape narrative.

Figure 3 further highlights certain criteria was highly present in non-credible statements; descriptions of interactions (71.4%) and reproduction of conversation (71%), and false statements; quality of detail (85.7%), unusual details (85.7%) and admitting lack of memory (71.4%).

**Table 2 [Apologies / ThisTable is not available]**

**Percentage presence of criteria present in Genuine and Credible Statements**

**Inter-rater reliability.**

To assess reliability of the investigators scoring, a random sample (N=8) was also independently rated by an external investigator. A Pearson's correlation was computed (r=.98, p<0.001). This score and figure 3 (below) indicates there was a strong positive significant correlation between the investigators and independent investigators rating, indicating high reliability.

![Figure 3: Demonstrating the relationship between the investigators and independent CBCA Scoring](image)

**Additional Phenomena**
A graph displaying the means and standard deviations for each type of statement veracity was created (see figure 4). Figure 4 depicts to what extent each statement score may overlap with another type of veracity. It can be seen from figure 4, there is an expected overlap of scores between genuine (M=18.86, SD= 3.07) and credible statements (M=16.57, SD=3.60), and also between false (M=6.29, SD=1.38) and non-credible statements (M=6, SD=2). However figure 4 also depicts there is a possible overlap between the scores for credible statements (M=16.57, SD=3.60) and scores for non-credible statements (M=6, SD=2).

Discussion

Main Findings

These results support the investigations hypotheses. Hypothesis one can be accepted as the findings demonstrated a significant difference between the CBCA scores of genuine statements, and false statements. The second hypothesis can also be accepted as the findings also depicted a significant difference between the CBCA scores for credible statements and non-credible statements. Furthermore hypotheses three and four can be accepted, as no significant difference was observed between the CBCA scores of genuine and credible statements or false and non-credible statements. It can be argued the main aims of the investigation have been achieved and the CBCA scale has demonstrated an ability to assess the complexity of genuine and credible adult rape narrative.
Links to Research

The results also support the findings of the previous research in showing CBCA as an appropriate methodology for assessing statement veracity. Most importantly these findings support the Undeutsh (1982; 1984) hypothesis, as they conform to the view a genuine statement will differ in content, quality and characteristics from a false allegation and supports the work of Holsti(1969), who claimed inferences can be made about one’s psychological state through systematically identifying specific criteria in statements.

The findings of this investigation are also concurrent with the findings of previous CBCA field investigations (Esplin et al. 1988; Boychuk, 1991; Lamb, Sternberg & Esplin et al., 1997; Hershkowitz, Lamb & Sternberg et al. 1997) in which genuine statements receive higher scores than false allegations. Moreover the findings also indicate, as per previous research suggestion, that the CBCA is not exclusive to child statements, and can be applied to adult narrative (Kohnken, 1995; Porter & Yuille, 1996; Ruby & Brigham, 1997; Steller & Kohnken, 1989) and most significantly these findings also reflect the findings of Parker & Brown (2000), advocating the use of CBCA to accurately differentiate between a genuine and false statement of adult rape.

The findings illustrating the presence of criteria in statements are consistent with Vrij’s (2005) review of field CBCA studies. Vrij (2005) illustrates logical structure, quantity of details, contextual embedding and details characteristic of offence that were most present in field CBCA research, findings concurrent with the results of this investigation. Moreover Vrij (2005) illustrates in previous research accurately reported details misunderstood, raising doubts about one’s testimony and self-depreciation were those least present. Again the findings of this investigation support this, with the exception of attribution of the perpetrator's mental state, present in 50% of the statements. Furthermore this investigation found 0% presence of pardoning a perpetrator, a contrast to the findings of Vrij (2005). This however could be explained by the nature of adult rape narrative, as a victim is unlikely to pardon a perpetrator or admit doubts in their testimony to an investigating officer, through fear of affecting the allegations credibility.

The criteria found highly present in false statements (quantity of detail, unusual detail, admitting lack of memory) and non-credible statements (descriptions of interactions, reproduction of conversations) is not unexpected, as it is likely an adult making a false allegation would include such factors in order to make their statement ‘ring true’.

Therefore it could be argued revisions of the CBCA are required to adapt it fully to rape narrative.

Additional Phenomena

The standard deviations illustrated an over-lap between the scores of credible and non-credible statements. This could be significant for any application of the scale to police practice, as it illustrates a certain percentage of statements rated credible may in fact be non-credible. Conversely this phenomenon may be
explained as a result of the limited sample used, and a larger sample may result in narrower standard deviations between scores.

Benefits

These findings may have implications beyond supporting the CBCA, in particular in relation to police investigations. As mentioned earlier, aside from the CI (Geiselman & Fisher, 1989) the police have no standardized procedure to assess the veracity of statements, relying on ‘short-hand cues’ (Marshall & Alison, 2006). However research has suggested officer’s ability to detect deception in these manners is no better than chance (Ekman & O’Sullivan, 1991). Therefore, although unable to offer 100% accuracy, it could be argued the CBCA scale (or revised form) could be utilized as a useful investigative tool in order for police to gain a systematic indicator of the possible veracity of a rape statement, which could be significant in future questioning or opening new lines of enquiry.

Limitations

There are several limitations which must be addressed. The first of these limitations, also pointed out by Vrij (2005) and Parker & Brown (2000), is the majority of the sample was only scored by a single investigator, and only a small sample (N=8) was scored by an independent investigator. There it could be argued the data may be subject to a degree of investigator bias.

A second methodological issue is the small sample size. Rape is a very distressing and sensitive issue, and access to real statements is restricted. Therefore due to time and authorities constraints only a small sample (N=28) could be utilized. It is worth noting the majority of investigations tackling issues associated with rape also use small samples, therefore this may be considered the norm. However this doesn’t detract from the notion that, with greater time and access available to statements, a wider range of statements could have been analysed. Furthermore researchers have argued CBCA scores are affected by several extraneous variables and individual differences (Buck, Warrern, Betman et al., 2002; Lamers-Winkleman & Buffing, 1996) and a larger sample may eradicate the chances of the data being affected.

A third methodological issue of CBCA is the limited number of channels analysed. CBCA only focuses upon what is spoken, recorded and/or transcribed from interviews; however it is important not to ignore the significance of non-verbal behaviour and non-verbal cues on detecting deception. Research has empirically illustrated the benefits of non-verbal behaviour such as hand movements (Vrij, 1997) and body language (DePaulo, 1994; Zuckerman & Driver, 1985; DePaulo, Stone & Lassiter, 1985) to detect deception. Moreover in the current technological age there is an increasing amount of technology which can detect deception using multiple channels, such as computer facial analysis (Pantic & Rothkrantz, 2000; Lanitis, Taylor & Coates, 1997), micro-expressions (Bartlett et al., 1999), FRMI (O’Craven & Kanwisher, 2000), PET scans (Farwell & Smith 2001) and artificial neural networks (Rothwell et al., 2006). These advancements suggest time-consuming processes, like CBCA, may soon become out-dated.
A forth limitation is that there are scenarios of false allegations it is impossible to detect with CBCA. The first scenario is when an individual has been raped, but identifies the incorrect person as the perpetrator (Vrij, 2005). A statement of this type will embody all the criteria of a genuine statement (obtaining a high CBCA score); however it will remain a false allegation against the innocent individual. The second scenario of false allegation CBCA cannot detect is when the rape has been imagined or a situation misunderstood. Additional research has illustrated imagined narratives can be very coherent and detailed (Ceci, Huffman, Smith, & Loftus, 1994; Ceci, Loftus, Leichtman, & Bruck, 1994; Porter, Yuille, & Lehman, 1999), and subsequently would obtain a CBCA scale, indicating authenticity. The final scenario suggests well prepared coerced answers can also be very internally consistent, and individuals can be taught to construct statements which feature many of the criteria on the CBCA scale and would be classified as genuine (Joffe & Yuille, 1992; Vrij, Kneller, & Mann, 2000).

A fifth methodological issue is the level of familiarity of the investigator with the CBCA scale. Researchers have argued different levels of training are required to become a CBCA expert, ranging between 2-3 day work-shops (Akehurst, Vrij & Bull, 1996; Landry & Brigham, 1992) to three week courses (Kohnken, 1999). However findings into training effects are mixed with certain researchers finding significant effects (Landry & Brigham, 1992), no effect (Ruby & Brigham, 1998; Santilla et al., 2000) and some training even made performance worse (Akehurst et al., 1998; Kohnken, 1987). However the high correlation between the investigators and independent investigators scores reflect Vrij’s (2005) meta-analysis, suggesting the investigator unfamiliarity had no effect on the score.

One final criticism comes from feminist theorists. The feminist perspective outlines alleged false allegations of rape are simply cases of rape which are not believed by investigating officers (Russell, 1984). Therefore the sample groups for this investigation, albeit established through ground-truths, will still court criticism from the feminist perspective. However it is worth noting this study is certainly not alone in indicating false allegation of rape is a very real issue (Theilade & Thompsen, 1986; Kanin, 1984).

Future Research

A logical expansion of this study would expand the sample and focus on additional individual factors, such as age, gender, suggestibility, alcohol/drug consumption or interview style (Vrij, 2005). Research has indicated age can be significant in CBCA scores (Anson et al., 1993; Boychuk, 1991; Craig et al., 1999; Vrij et al., 2002) for example Buck et al. (2002) found age and total CBCA scores were positively correlated. Looking at additional factors was not possible in this investigation due to time and authority constraints, however investigation of a larger sample may affirm or refute the use of CBCA as an investigative tool. Furthermore the application of a validity check-list (Vrij, 2005; Raskin & Esplin, 1991) could be beneficial. Various validity check-lists have been proposed, considering elements such as; mental capacity, susceptibility, motives, coercive questioning and consistency of other evidence amongst other criteria (Vrij, 2005).
A second extension would be to have trained police officers, independent investigators and external judges conduct CBCA on a series of rape statements, and assess whether the scores assigned are significantly different or correlated. This would be beneficial as research has posited police officers are no better at detecting deception than lay persons (Ekman & O'Sullivan, 1991; Ekman, O'Sullivan, & Frank, 1999), or police officers may be more sceptical than independent researchers (Ruby & Brigham, 1997). Furthermore this extension could incorporate a qualitative remark from the police officers and independent judges as to why they have assigned the chosen mark. This would be significant in exploring key differences in decision making, expertise and rationales between police officers, investigators and external judges, and whether police training influences an officer’s ability to detect deception. This could suggest whether the CBCA analysis could be an investigative police tool, or whether police officers judgements on the CBCA scale are biased or sceptical.

Furthermore a third beneficial study would tackle the possibility of an overlap between credible scores and non-credible scores. If a larger sample still demonstrated an overlap between credible and non-credible statements, it would be beneficial to conduct a study which establishes confidence bands, to ascertain the likelihood that a credible statement may in fact be non-credible. This would be essential for application in police practice.

A fourth extension could be to develop a revised form of the CBCA scale. As outlined earlier certain criteria didn’t appear in many or any of the statements analysed, as they are likely to diminish the credibility of the allegation. Therefore it may be beneficial to develop and test a revised form which eliminates these criteria, replacing them with criteria which appears more prominent in rape narrative such as ‘reference to physical harm’ or ‘reference to offenders clothing’.

A fifth extension could combine CBCA alongside a measurement of one of the other multitude of different channels cited in deception literature. For example police officer judgement of body language (Ekman & O’Sullivan, 1991) thermal imaging (Warmelink et al. 2011) or facial analysis (Pantic & Rothkrantz, 2000). Marshall and Alison (2006) illustrated that controlled observation of behaviour in interviews can be indicative of detecting false allegations of rape. Findings showed genuine statements reported a larger total of behaviours than simulated interviews. Furthermore pseudo-intimate behaviours were reported more in genuine statements, whereas violent behaviour is often more present in simulated interviews, which appear more behaviourally incoherent than genuine statements (Marshall & Alison, 2006). Therefore combing an additional measure, alongside CBCA, may create a more accurate methodology of detecting deception.

**Conclusion**

To summarize, this investigation strengthens the application of CBCA to test veracity of adult statements, and illustrates potential application for investigative policing. There are many possible expansions and revisions of the CBCA, with potential implications for policing, and therefore it is important that research in this area continues.
References


Institute on the Child Witness in Court: Cognitive, Social and Legal Perspectives, Lucca, Italy.


