Attributes of Truthful Versus Deceitful Statements in the Evaluation of Accused Child Molesters

Shawn Johnston  
*Portland State University*, shawnjohnstonphd@comcast.net

Alexis Candelier  
*Portland State University*

Dana Powers-Green  
*Portland State University*

Syeda Rahmani  
*Portland State University*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/psy_fac](https://pdxscholar.library.pdx.edu/psy_fac)

Let us know how access to this document benefits you.

**Citation Details**

Johnston, S., Candelier, A., Powers-Green, D., and Rahmani, S. (2014). Attributes of Truthful Versus Deceitful Statements in the Evaluation of Accused Child Molesters. SAGE Open. DOI: 10.1177/2158244014548849

This Article is brought to you for free and open access. It has been accepted for inclusion in Psychology Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.
Attributes of Truthful Versus Deceitful Statements in the Evaluation of Accused Child Molesters

Shawn Johnston¹, Alexis Candelier¹, Dana Powers-Green¹, and Syeda Rahmani¹

Abstract
The ability to detect deception, in everyday social interactions and psychological evaluations, can literally mean the difference between life and death. Beyond physiological and nonverbal techniques for detecting deception, research has focused on criteria designed to evaluate the content of verbal statements to distinguish between true or actually experienced events versus internally manufactured or fabricated events. Criteria from two techniques that have received empirical support, criteria-based content analysis and reality monitoring, were used to create an 11-item Deception Detection Checklist (DDCL). In this study, 130 college undergraduates used the DDCL to rate the exculpatory statements of two accused child molesters: one truthful, the other untruthful. The 11 items composing the DDCL, as well as a measure of perceived truthfulness, were all scored on 7-point Likert-type scales. Nine of the 11 items on the DDCL significantly differentiated between the true and untrue statements in the predicted direction. Overall scores on the DDCL indicated that the false statement was rated as significantly more deceptive than the true statement. The DDCL possessed good reliability, and a series of factor analyses provided strong support for the construct validity of the measure. The 7 psychometrically strongest items from the DDCL included variables assessing the extent to which statements included clarity of detail, spatial details, temporal details, and contextual details, as well as the relevance, reconstructability, and realism of the statement. These results indicate that subjects were able to use this measure to reliably differentiate between true and false statements made by accused child molesters.

Keywords
detecting deception, clinical interviewing, forensic psychology

In a recently published book, ten Brinke and Porter (2013) describe considerable research indicating that people in general are rather bad at detecting lies, with most rating little better than chance at detecting deception. Distressingly, research also indicates that police officers and trial court judges, those society has entrusted with the job of dispensing justice, may be no better at detecting deception than the average person. For example, research relates that police officers are actually trained to focus on signs of nervousness, statistically one of the weakest predictors of deception. In their published work, ten Brinke and Porter cite research indicating that judges and jurors, as well as people in general, are actually less likely to attribute deceit to attractive rather than unattractive persons. In all, this research leads us to believe that people are just lousy at catching liars and at least part of the reason is that most of us focus on the wrong overt behaviors.

While some lies can be innocuous, even benign, depending on the vulnerability of the person and the seriousness of the situation, some lies can have devastating even deadly results. This is obviously true in police investigations as well as criminal trials. It is also true, however, that a forensic evaluation, on which the court relies to make its decisions, can be completely wrong in the assessment of client dangerousness and thereby have terrible consequences for the community into which the client/defendant is returned. It is no wonder that the perceived sincerity or honesty of other people is often indicated as the single most desirable personality trait. The ability to correctly identify truthful versus deceitful persons represents an important social-cognitive skill with seemingly obvious adaptive implications. Within the context of judicial proceedings or court-ordered forensic psychological evaluations, this is doubly true.

¹Portland State University, OR, USA

Corresponding Author:
Shawn Adair Johnston, Portland State University, 1825 SW Broadway, P.O. Box 751, Portland, OR 97207, USA.
Email: shawnjohnstonphd@comcast.net

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 3.0 License (http://creativecommons.org/licenses/by/3.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (http://www.uk.sagepub.com/aboutus/openaccess.htm).
**Literature Review**

Most research on deception detection has focused on the assessment of physiological or nonverbal responses, commonly referred to as body language. For example, the fact that many people exhibit autonomic nervous system responses when they lie is the basis of the polygraph examination and the guilty knowledge test. Even more common than that, probably every one of us has had the experience of observing another person begin to sweat and breathe more rapidly when he or she has begun to spin a lie. Sweating, breathing heavily, excessive blinking, avoiding eye contact, and general physical agitation all at least have some empirical support indicating that such behaviors, with some people, in some situations, can indicate the attempt to deceive. While physiological responses to telling lies is a well-documented phenomenon, such responses have also been shown to occur in reaction to embarrassment and other powerful emotions, having nothing at all to do with deception.

Another approach to detecting deception derives from the idea that true statements can be differentiated from false statements as a function of the cognitive task or challenge represented by attempting to deceive another person, a general theory known as content complexity (ten Brinke & Porter, 2013; Trivers, 2011) or cognitive load (Vrij, Mann, & Fisher, 2006). From this theoretical perspective, the liar has taken on a complicated and weighty cognitive chore by attempting to deceive others. Not only must the liar make certain to suppress his own knowledge of the truth, he also needs to be relatively certain that the target of his deception is unaware of that truth. Beyond this, the liar is also required to be on constant guard against making inconsistent or contradictory statements that could reveal his deception. This overextension of cognitive ability appears to be the reason why liars tend to make shorter statements than truth tellers, as well as why they are more likely to pause excessively while verbalizing their statements. We believe that the cognitive load experienced by a deceitful accused child molester would be especially challenging and they would exhibit multiple signs of this attempt to deceive. While there is some empirical support for techniques using physiological and nonverbal cues to differentiate between truthful and deceitful statements, our review of the deception detection literature suggests that systems based on verbal cues or information may be more reliable.

With regard to the assessment of deception using verbal cues, two techniques have been independently developed and tested for differentiating between truthful and deceitful statements. The first of these, criteria-based content analysis (CBCA), was originally developed in Germany for assessing the veracity of children claiming to have been sexually abused (Roma, San Martini, Sabatello, Tatarelli, & Ferracuti, 2011; Vrij, 2008). The second technique, reality monitoring (RM), was developed to help differentiate between the actual and hallucinated experiences of schizophrenic individuals (Johnson & Raye, 1981; Vrij, 2008). While the original goals and foci of these two techniques could hardly have been more different, rather interestingly, both of them fundamentally revolve around the idea that descriptions of actual versus imagined or fabricated experiences will differ in predictable ways. Events that we literally experience in the real world will have significant external inputs, while events generated or created in the imagination or mind’s eye will have primarily, if not exclusively, internal inputs. Both techniques predict then that actual or real experiences versus imagined or fabricated ones, when described verbally will contain more and richer details, appear more plausible and realistic, be imbedded in time, space and some social or interpersonal context, and include perceptual and/or sensory information, such as the speaker’s experience of temperature, sound, smell, and taste. Johnston et al. (2013) reviewed in detail the published research regarding specific criteria from both techniques, which had been subjected to empirical study with regard to their ability to differentiate between truthful and untruthful statements, and found that both systems demonstrated the ability to differentiate between truthful and untruthful statements to a significant degree (Vrij, 2008).

From the 19 criteria used in the CBCA and the 8 criteria used in RM to distinguish truthful from untruthful statements, 11 items were derived, which are believed to capture the most important features, in particular, those concepts that appear in both systems (see Figure 1). Item 1 assesses the clarity of detail provided in a statement. Item 2 assesses the presence of perceptual information (e.g., the experience of temperature or taste). Item 3 assesses the presence of spatial information, such as physical proximity of people and objects. Item 4 assesses the presence of temporal information, such as how events are connected in time. Item 5 assesses the expression of affect or emotion on the part of the person making the statement. Item 6 assesses the reconstructability of the statement, particularly whether the information provided is consistent, noncontradictory, and logically plausible. Item 7 assesses how realistic the statement is. Item 8 assesses how self-serving the statement is. Item 9, which is our manipulation check variable, not a part of the checklist per se, assesses the overall perceived truthfulness of the statement. Item 10 assesses the extent to which information in the statement is embedded in an interpersonal or social context. Item 11 assesses whether statements made by other individuals are reproduced versus being paraphrased or inferred by the person making the statement. And finally, Item 12 assesses how relevant are the details provided in the statement. Unlike the dichotomous rating systems used in CBCA and RM, the 11 new items were designed to be rated on a 7-point Likert-type scale, going from **strongly disagree** to **strongly agree**, that the attribute, such as clarity of detail, was present in the statement. The individual response scales were expanded from 2 to 7 options to provide research subjects with more finely differentiated alternatives and to
increase the reliability of the measure. It should also be noted that all items are scored in the direction of their name, that is, a higher score on the spatial information item means that more spatial information was present in the statement.

**Method**

**Independent Variable**

The independent variable in this study consisted of two conditions: an exculpatory statement made by an accused “truthful” child molester and an exculpatory statement made by an accused “untruthful” child molester. Both statements were taken from the files of the senior author who has specialized in the evaluation and treatment of accused and convicted sexual offenders since 1980. The two alleged child molesters were both referred for confidential psychological assessment by their attorneys regarding issues such as sexual dangerousness and probation suitability. The client statements focused exclusively on the client’s explanations of the charges against him and why he believes that he had been falsely accused of numerous sexual offenses against multiple children. Both clients’ statements were taken directly from the relevant clinical interview notes for that client and were comparable in length, clear, and relatively articulate. Beyond this, the single most important reason for selecting these particular client statements was the actual judicial outcomes in each of their respective criminal cases. Here, it should be noted that because it is impossible to absolutely know when a statement is true or false, it is often the case in the deception detection literature that a client statement is viewed as truthful when a judge or jury has exonerated the client. Correspondingly, a client statement is viewed as false when a judge or jury finds the client guilty. Thus, the statement identified as truthful in this research was made by an individual who was acquitted on all counts by a jury in less than 2 hr, with numerous jurors reported as saying that the defendant was obviously innocent rather than simply not guilty. In contrast, the untruthful statement was made by an individual who was convicted on all counts by a jury within a comparable period of time. While this is obviously an imperfect method for selecting truthful versus untruthful statements, it is widely accepted in the research literature concerning deception detection and does contain the virtue of involving real people, accused of very serious crimes, which were carefully evaluated by juries of their peers.

In addition, three practicing forensic psychologists, with approximately 90 years of experience between them, were asked to review the statement of the untruthful accused child molester. They were then asked whether they found the statement truthful or untruthful and, if they believed that the statement was untruthful, to rate how typical it was of untrue

| Reality Monitoring | Deception Detection Check List | Content Based Criterion Analysis |
|-------------------|-------------------------------|---------------------------------|
| Clarity           | Situational Attributes        | Logical Structure               |
| Perceptual Information | Clarity of Details     | Reproduction of Conversation    |
| Spatial Information | Reproduction of Conversation | Logical Structure               |
| Temporal Information | Reproduction of Conversation | Reproduction of Conversation    |
| Affect            | Judgement Attributes          | Accounts of Subject’s Mental State |
| Reconstructability of Story | Reconstructability of Story | Descriptions of Interactions    |
| Realism           | Realism                       | Details Characteristics of the Offense |
| Not Included Items | Self-Serving Statements       | Raising Doubts About One’s Statements |
| Cognitive Operation | Relevance of Details          | Self Depreciations              |
|                   |                               | Pardoning the Perpetrator       |
|                   |                               | Superfluous Details             |
|                   |                               | Not Included Items              |
|                   |                               | Accurately Reported Details     |
|                   |                               | Misunderstood                   |
|                   |                               | Admitting Lack of Memory        |
|                   |                               | Attributions of Perpetrators    |
|                   |                               | Mental State                    |
|                   |                               | Related External Associations   |
|                   |                               | Spontaneous Corrections         |
|                   |                               | Unexpected Complications during Incident |
|                   |                               | Unstructured Production         |

**Figure 1.** Summary of content of items/criteria composing reality monitoring, Deception Detection Checklist, and criteria-based content analysis.
statements by accused and convicted child molesters. All three psychologists indicated they believed that the statement was untrue and that the statement was, in their experience, typical of statements made by untruthful child molesters.

**Dependent Variable**

As indicated above, 11 items were designed, inspired by the research literature from CBCA and RM, to assess different attributes or dimensions indicative of truthful versus untruthful statements. The content of these items and their origins is discussed in the literature review. We hypothesized that the research subjects would rate the truthful statement significantly higher on the items believed to be correlated with truthfulness, that is, clarity of detail, presence of perceptual/sensory information, realism of the statement, and so on. However, for the item assessing the self-serving nature of the statement, higher scores were hypothesized to be indicative of greater deceitfulness.

The 11 items were then combined into a Deception Detection Checklist (DDCL). The scores on the 10 DDCL items indicating truthfulness had to be reversed, so overall DDCL score would indicate greater deceptiveness. Overall scores on the DDCL can vary from a low of 11, the number of items, to a high of 77, seven times 11, which would indicate the maximum possible deceptiveness on the scale. It should be noted that a 12th item was included on the checklist though it does not contribute to the overall checklist scores. This item, also on a 7-point response scale, asks the subject to rate the truthfulness of the two different statements. Scores on this item, in part, represent an important check on how successful we were in manipulating the independent variable. We defined successful manipulation of the independent variable, as subjects correctly rating the truthful statement as truthful and the untruthful statement as untruthful.

**Participants**

Subjects were 130 students from a large university in Oregon enrolled in an undergraduate class in forensic psychology.

**Procedures**

During the last week of class, the student subjects in this research were briefly described the 11 items from the DDCL and the research from which they originate. As noted above, these were students in an upper division forensic psychology class and much of the school term had been given over to discussion of sociopaths, sex offenders, and the reliable psychological assessment of such client populations. While the topic of deception had been touched on many times before in this class, with regard, for example, to its threat to psychological testing and interviewing, this was the first discussion of the specific 11 variables previous research suggested are capable of differentiating between truthful and untruthful statements. Approximately 10 min were necessary to explain the DDCL items and the rating task with which the subjects were about to be presented.

At this point, the questionnaire packets were handed out to the subjects. The questionnaire first informed the subject that their involvement in this research was entirely voluntary, would have no effect on their grade in the class, and was, in fact, completely confidential and anonymous; that is, no subject was asked for personal identifying information other than age, gender, and major.

Subjects were then asked to read both statements and then rate the two statements relative to the items on the checklist, also provided to the subjects. Thus, all subjects who participated in the research, which appeared to be about 100% of the members of the classes, rated both statements per the questionnaire items. Subjects required approximately 15 min to read the statements and complete the questionnaire. Upon completion of the task, approximately half an hour was given over to explaining the background and hypotheses of the research while also soliciting subject reaction to the questionnaire, the rating task, and the goals of the research. Overwhelmingly, the subject’s responses to the task were positive, indicating that the task was interesting and useful.

**Results**

The internal consistency reliability coefficient (Cronbach’s alpha) for the DDCL was computed at .80. The means and standard deviations for the truthful and untruthful conditions along with the t-test statistics for the DDCL, the 11 items that compose it, and the one manipulation check variable are presented in Table 1. We would first bring the reader’s attention to the results regarding the manipulation check variable, as these results pertain to what extent we were successful in having the subjects perceive the truthful statement as true and the untruthful statement as untrue. If the subjects failed to perceive any difference in truthfulness between the true and untrue statements, it would be difficult to argue that scores on the DDCL differed as a function of the statements being true or untrue. To the extent that the subjects perceived a significant difference between the test statements in truthfulness, differences observed between scores on the other 11 items can reasonably be assumed to be directly connected to the overall perception of the statement’s truthfulness.

The results of the manipulation check variable indicate a very significant difference in the perceived truthfulness of the truthful versus untruthful statements. Indeed of all the t-tests performed, the largest t-test statistic and greatest degree of significant difference was found between subject’s overall ratings of the truthfulness of the truthful versus untruthful statements. These results indicate our success in providing two different statements to subjects, which they readily
Johnston et al.

5

distinguished as truthful versus untruthful (see Table 1 for t-test statistics).

Next, we would like to bring the reader’s attention to the pattern of means and standard deviations presented in Table 1. As indicated above, all the 11 items composing the DDCL as well as the manipulation check item were rated by subjects on 7-point scales. Whether the individual item scales were unipolar, going from none of the attributes to a maximum of it, or bipolar, going from the extreme of untruthfulness to its opposite of total truthfulness, the score of 4 represents the midpoint of every item scale. Thus, mean scores below 4 would indicate untruthfulness, whereas mean scores above 4 would indicate truthfulness. The potential importance of this point is that it permits us to indicate not only whether an item is able to differentiate between levels of truthfulness but also whether that item is able to differentiate between truthfulness and untruthfulness per se. We believe that it is important for the mean scores on individual items in the truthful and untruthful conditions to fall on opposite sides of their respective scale midpoints because we are interested in not just the assessment of truthfulness along a continuum but categorically as well.

With regard to the 11 items composing the DDCL, subject’s ratings differed significantly in the predicted direction in all but 2 of the items (expression of affect and self-serving statements). Importantly, 7 of the 11 items exhibited both significantly different ratings between conditions as well as containing mean scores for the untruthful statement on opposite sides of the item scale midpoint than the mean scores for the truthful statement. These seven variables included clarity of detail, spatial information, temporal information, reconstructability of the statement, realism of the statement, contextual information, and relevance of the details provided in the statement. In addition, 2 of the items, perceptual detail and reproduction of conversation, yielded significant differences between the truthful and untruthful statements, but the mean scores for both conditions were on the same side of the item scale midpoints.

In addition to examining the ability of the items to differentiate between the true and untrue statements, we thought it would be interesting to see how each of the 11 items correlated with both the DDCL overall score and our measure of perceived truthfulness. We assumed that some of the items, especially those assessing the realism of the statement, would be more strongly associated with perceived truthfulness than other items such as temporal or spatial details. The perception of how realistic a statement is appears to go to the heart of whether it is perceived as truthful. In addition, we were also interested in the relationship between the DDCL overall score and perceived truthfulness, in that, both these variables represent different approaches to the assessment of truthfulness: one directly, by simply asking the subject how truthful they believed the statement to be, and the other indirectly, by computing an overall score composed of items believed to be characteristic of truthfulness. Table 2 presents the correlations between the 11 items and both the DDCL score and the measure of perceived truthfulness. In addition, the direct correlation between DDCL scores and perceived truthfulness is also reported in Table 2.

For ease of reading Table 2, the perceived truthfulness variable was scored in the negative direction, thus obviating the need for plus or minus signs. With this in mind, the correlation between perceived truthfulness and the DDCL is rather strong (.57). Put another way, almost one third of the variance in the perceived truthfulness of the statements is accounted for by DDCL scores. Not surprisingly, the

| Table 1. Means, Standard Deviations, and t-Test Values for the Truthful Versus Untruthful Statement on the Measure of Perceived Truthfulness, DDCL Scores, and the 11 Items Composing the DDCL. |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                | Truthful statement | Untruthful statement | t tests | p |
|                                | M   | SD  | M   | SD  | M   | SD  | t tests | p  |
| Perceived truthfulness         | 4.93| 1.33| 2.47| 1.26| 12.43| <.001|
| DDCL scores                    | 38.93| 10.31| 52.98| 11.47| 9.66| <.001|
| Clarity                        | 5.10| 1.44| 3.49| 1.71| 7.58| <.001|
| Perceptual information         | 2.91| 1.56| 2.41| 1.24| 3.22| .001|
| Spatial information            | 5.12| 1.55| 3.18| 1.62| 9.59| <.001|
| Temporal information           | 5.12| 1.45| 3.52| 1.69| 7.88| <.001|
| Affect                         | 3.26| 1.76| 3.73| 1.84| 2.09| <.020|
| Reconstructability of story    | 4.19| 1.66| 2.77| 1.46| 6.98| <.001|
| Realism                        | 5.26| 1.34| 2.94| 1.42| 11.66| <.001|
| Self-serving statements        | 4.07| 1.72| 5.30| 1.94| 4.47| <.001|
| Contextual information         | 5.41| 1.24| 3.73| 1.58| 8.95| <.001|
| Reproduction of conversation   | 3.31| 1.70| 3.04| 1.78| 1.31| <.100|
| Relevance of details           | 5.46| 1.18| 3.51| 1.69| 10.31| <.001|

Note. DDCL = Deception Detection Checklist.
individual items of the DDCL are, with one exception, more strongly correlated with overall DDCL scores than the perceived truthfulness variable. One would expect this given that the overall DDCL score contains within it each of the 11 items with which it is being correlated. One would also expect these correlations to be high if all the items from the DDCL do, in fact, measure different aspects or dimensions of truthfulness, an assumption to be specifically explored below. Also not surprising, given the t-test results, is the poor performance of the affect and self-serving variables with regard to their correlations with both the DDCL and perceived truthfulness. In all, for the most part, the DDCL items had correlations about twice as strong with the DDCL as with the measure of perceived truthfulness. All but one of these correlations were statistically significant, suggesting that both the direct and indirect approaches to assessing the truthfulness of a statement have utility.

With regard to our expectations regarding the possible differential pattern of correlations, we found it very interesting that the realism variable was by far more strongly correlated with perceived truthfulness relative to any other item from the DDCL. Indeed, the realism variable was the only one to be equally strongly correlated with both DDCL scores and perceived truthfulness ($r = .64$). In all, with two conspicuous exceptions, the items from the DDCL significantly differentiate different levels of truthfulness as well as reveal a pattern of correlations between the overall DDCL scores and perceived truthfulness, which is generally consistent with expectation as well as being consistent with our clinical and forensic experiences.

Next, we performed two series of factor analyses on our data. The first set of factor analyses involved the 9 variables capable of significantly differentiating between the true and untrue statements in the predicted direction. Thus, the expression of affect and self-serving statements variables were excluded from all further analyses. For the purpose of the first factor analyses, we entered all 9 variables from both the truthful and untruthful conditions, yielding a grand total of 18 variables on which the factor analysis was performed. There were a number of reasons we felt this would be the most appropriate approach for performing an exploratory factor analysis. Specifically, entering all 18 variables into the analysis affords us the opportunity to compare head on the relative importance of our variables with regard to their ability to differentiate between truthful and untruthful statements. For example, do all of the variables have approximately the same strength of relationship with truthfulness versus untruthfulness or is it possible that some variables are more closely associated with ratings of the true or the untrue statement? In this regard and with reference to the concept of factorial validity, we were hoping that our variables would yield a factor solution in which the factors clearly revolved around the themes of truthfulness and untruthfulness. In all, we believed that this factor analysis would speak directly to the construct validity of our variables as measures of both truthfulness and untruthfulness while suggesting possible differential contributions to the identification of the true versus untrue statements on the part of individual items.

The first factor analysis yielded three components. The first two factors clearly revolved around the themes of untruthfulness versus truthfulness, while the third factor accounted for a trivial amount of variance and was based almost entirely on the single item of reproduction of conversation. This was not especially surprising as this variable was one of the weakest of the 11 items in our statistical tests of significant differences between the conditions. Criteria associated with factor manipulation indicated the appropriateness of imposing a two-factor solution on our data (results are presented in Table 3).

It should first be noted that every variable used to assess the untruthful statement loaded on Factor 1, which for obvious reason we have chosen to call the Untruthful factor.
Similarly, every one of the variables used to assess the truthful statement loaded on Factor 2, which we just as obviously will refer to as the Truthful factor. As there are no significant loadings on Factor 1 of items used to rate the true statement, and no significant loadings on Factor 2 of items used to rate the untrue statement (in fact, all of these factor loadings hovered around 0), we have left these data out of the table.

We believe that these results are extremely interesting. First, and as already alluded to, the two-factor solution consists of Factor 1 containing only variables used to rate the untrue statement and Factor 2 containing only variables used to rate the truthful statement. We have inferred then that the underlying theme of Factor 1 is untruthfulness and the underlying theme of Factor 2 truthfulness. We also believe that it is interesting to note the “purity” of the factors in that the nine variables contributing to the analysis are very different in content. While spatial and temporal information intuitively seems connected, just like realism and reconstructability, the first two variables do not appear to have any obvious connection with the latter two. In other words, it seems clear from the factor analytic findings that the most powerful underlying theme in the data is in fact the truthfulness versus untruthfulness of the statements rated by our subjects.

It also bears noting that with one exception, the reproduction of conversation variable in the untruthful condition, every factor loading is between 0.6 and 0.8. While this is not surprising as in the three-factor solution, reproduction of conversation essentially generated and accounted for the third factor, it does suggest that otherwise all of our variables contribute more or less about the same to the respective factors on which they load. Indeed with the one exception of the reproduction of conversation variable, not one of the other eight variables differed more than 0.1 from the truthful to untruthful conditions. In other words, no obvious difference was observed in any of these eight variables whether they were being used to assess the true or untrue statement. With the exception of the reproduction of conversation variable, not one other variable appears to have been differentially connected to the factor on which it loaded. Thus, clarity of information was approximately as important in assessing an untruthful statement as it was in assessing a truthful statement, and none of these eight variables was conspicuously more important in its association with truthfulness versus untruthfulness. While the eight variables differ with regard to their factor loadings, going from 0.6 to 0.8, a difference which itself represents a comparatively small range, the connection between the individual variables and the factors on which they load is roughly equivalent. In terms of the concept of factorial validity, these data appear particularly encouraging.

Our final series of analyses involved performing two factor analyses: one on the items used to rate the truthful statement and the other on the items used to rate the untruthful statement. For these two analyses, we used only those items that were found to be significantly different in the predicted direction and whose item means fell on opposite sides of the scale midpoints in their ratings of the truthful versus untruthful statements. Thus, each of these seven items were capable

| Table 3. Factor Loadings for the 11 Items Composing the DDCL in the Combined Truthful and Untruthful Statement Conditions. |
|---------------------------------------------------------------|
| **Factor Loadings**                                          |
| **Factor 1**        | **Factor 2** |
|---------------------|--------------|
| **Untruthful statement**                                   |
| Reconstructability of story          | 0.771        |
| Reproduction of conversation         | 0.753        |
| Temporal information                  | 0.741        |
| Clarity                                 | 0.725        |
| Spatial information                    | 0.715        |
| Contextual information                | 0.701        |
| Perceptual information                | 0.684        |
| Relevance of details                  | 0.662        |
| Realism                                | 0.626        |
| **Truthful statement**                    |
| Temporal information                  | 0.800        |
| Spatial information                   | 0.772        |
| Reconstructability of story           | 0.761        |
| Contextual information                | 0.733        |
| Clarity                                | 0.718        |
| Perceptual information                | 0.667        |
| Realism                               | 0.648        |
| Relevance of details                  | 0.609        |
| Reproduction of conversation          | 0.423        |

Note. DDCL = Deception Detection Checklist.
of significantly distinguishing the true from untrue statement and did so in such a way as to indicate that the untruthful statement was false and the truthful statement was true. As discussed earlier, these seven items could be considered, psychometrically, our best variables as they not only distinguish different levels of truthfulness but can actually be used to distinguish between untruthful and truthful statements, per se. Beyond this, we also believed that conducting a factor analysis using these items in the truthful versus untruthful statement conditions would provide us with a direct opportunity to explore the possibility that the different variables may have differential relationships or connections with true versus untrue statements. While the results reported in Table 3 indicate that these seven variables are more or less equally associated with the assessment of truthful versus untruthful statements, it is definitely possible that the seven different variables may be relatively stronger or weaker in rating the truthfulness of a statement depending on whether the statement is true or false. In other words, it seemed likely that our variables would be differentially sensitive or important with regard to the dimension of truthfulness versus untruthfulness.

For example, as described above we believed that the variable of realism would be especially important in assessing whether a statement is true or untrue. By performing two factor analyses, one based on the variables used to rate the truthful statement and the other based on variables used to rate the untruthful statement, this would seem to afford the most direct opportunity to compare head on the importance of our seven best variables in assessing true versus false statements (results are presented in Table 4). The left side of Table 4 presents the results of the factor analysis on the items rating the truthful statement and the right side of the table presents the results of the factor analysis on the items rating the untruthful statement. We should note here that we expected this factor analysis would reveal some different relationships between the items and anticipated factors. Specifically, we reasoned that the variables connected with situational information, such as the spatial, temporal, contextual, and clarity of details, might generate their own factor given what seems to be the obvious conceptual connection between these items. Similarly, we reasoned that the reconstructability, realism and relevance of details, variables would be likely to form their own factor as these items also appear connected by a value judgment concerning the quality of the information in the statement. Candidly, we also expected that these latter three variables, especially realism and reconstructability, would prove relatively stronger than the situational variables in explaining the extracted factors as these three variables reflected judgments of the statement that seemed to directly overlap with the concept of truthfulness. Thus, while the presence of situational attributes is important in determining whether someone is speaking the truth, we reasoned that realism, reconstructability, and the relevance of what was said would be of primary importance in identifying truthful versus untruthful statements. We were, therefore, frankly surprised, albeit pleasantly so, to see that the two factor analyses yielded solutions composed of only one factor each. The factor analysis on the seven items used to rate the true statement created one factor only as did the factor analysis on the seven items used to rate the untrue statement.

Examination of the data presented in Table 4 indicates that our expectations regarding items differentially contributing to the factors received only very partial support. Indeed for the factor analysis on the truthful condition items, every one of the first four variables dealt with what we just described as situational details, while the three variables with the lowest loading were our so-called judgment variables. However, we do find it interesting to note that the single strongest item factor correlation for the untruthful condition is for the reconstructability variable even though realism and relevance remain more or less at the bottom of the list. It should of course be noted that the single-factor solutions for these two analyses speak directly to the construct validity of our measurement, indicating that each and every one of these seven variables does very well with regard to the assessment of both true and untrue statements. It also bears noting that the entire item factor loading in both analyses was relatively quite high with no loading falling below 0.6. Finally, it should be noted that more than 50% of the variance of the data was accounted for in both factor solutions. If we were going to be disappointed in our results, it is difficult to imagine a more pleasing disappointment.

| Table 4. Separate Seven-Item Factor Analyses for the Truthful and Untruthful Statement Conditions. |
|---------------------------------------------------------------|
| **Truthful statement** | Factor 1 | **Untruthful statement** | Factor 1 |
| Temporal information | 0.822 | Reconstructability of story | 0.780 |
| Spatial information | 0.791 | Temporal information | 0.776 |
| Contextual information | 0.772 | Contextual information | 0.731 |
| Clarity | 0.744 | Clarity | 0.712 |
| Reconstructability of story | 0.743 | Spatial information | 0.699 |
| Realism | 0.691 | Relevance of details | 0.687 |
| Relevance of details | 0.611 | Realism | 0.674 |

The table above shows the factor loadings for each variable in the truthful and untruthful statement conditions. The factor loadings range from 0.611 to 0.822, indicating a strong relationship between the variables and their respective factors. The factors are labeled as Factor 1 for both conditions, indicating a single-factor solution for each analysis.
Discussion

While the deception detection literature indicates that people are generally rather bad at catching liars, apparently most people assume that what others tell them is generally true, the subjects in our study appear to have exhibited a lie bias. Our results reveal that the untruthful statement was perceived as more deceptive than the truthful statement was perceived as true. In addition, the untruthful statement was rated as more deceptive with regard to overall DDCL scores than the truthful statement was rated as true on this same measure. A lie bias is hardly surprising, given that our subjects were upper division college students in a forensic psychology class assessing the statements of two men formally accused of child molestation. Even though it would seem desirable to have neither a truth nor lie bias, we have the sense that in forensic psychological evaluations where the stakes can include incarceration and/or the loss or gain of considerable amounts of money, a cautious, even wary, attitude on the part of the examiner may be a good thing. We also interpret these data positively in that they suggest, consistent with other research (ten Brinke & Porter, 2013), that empirically based techniques designed to detect deception can be taught.

Beyond this, we believe that the foremost implication of our results is in demonstrating that the truthfulness of a statement, at least in some situations, can be identified with a high level of reliability and internal validity. Clearly, the most conspicuous limitation of the present research is the fact that subjects rated only two statements, both pertaining to extreme and extraordinary situations. It will not be possible to make any serious claims regarding the external validity or generalizability of our measure until such time as it has been applied to the assessment of a greater variety of statements. Needless to say, the focus of future research would be to find out whether our measure, particularly the seven best items, is as reliable and significant in differentiating the truthfulness of client statements in a variety of other settings. While these items demonstrated the ability to differentiate between levels of truthfulness of accused sex offenders, it would be interesting to know whether they could do the same with regard to differentiating between clients actually suffering from rather than feigning mental illness.

While our subject’s ratings of the perceived truthfulness of the two statements were strongly associated with overall DDCL scores, the two measures are clearly not the same thing and approach the assessment of truthfulness in fundamentally different ways. Asking a subject to specifically rate the truthfulness of a statement represents as direct an assessment of truthfulness as one can imagine. However, of our seven psychometrically best items from the DDCL, four of them deal with aspects or attributes of a situation that have no a priori or necessary connection with truthfulness, per se. Why, for example, should information regarding spatial or temporal detail be so strongly correlated with the perceived truthfulness of a statement? While a clever sociopath could include in his or her statement these kinds of details, the robust nature of these variables speaks directly to the idea that true rather than internally manufactured experiences will include such “real world” information. Thus, while a knowledgeable sociopath could include such information in a false statement, it seems obvious that it would require considerable intellectual even scientific sophistication as well as substantial forethought and cognitive rehearsal.

In contrast, variables, such as relevance, realism, and reconstructability, which reflect judgments regarding a statement, seem to have a different relationship to the assessment of truthfulness, than situational attributes, such as spatial and temporal details. We believe that the relatively high correlation between realism and perceived truthfulness is illustrative of a uniquely important relationship between these variables. Of the many hundreds of psychological evaluations the senior author has performed on accused and convicted child molesters, it is striking how often the defendant’s version of the events which led to the criminal charges against him would seem unbelievable, if not outright absurd. It is entirely common for many of these men to claim that the primary reason they were being falsely accused of sexual abuse was that they were, in some important way, too nice, too generous, too caring, and too concerned regarding the victim and the victim’s family. While it is painfully apparent that innocent persons are at times falsely accused of serious crimes, in the real world, it is vanishingly rare for a child or children and their families to conspire together against a man, to falsely accuse him of molestation, just because that man had just been too kind or generous toward that family. The present results clearly suggest how critical to the assessment of truthfulness we believe the realism of the statement can be. It is our impression that when statements are conspicuously unrealistic, the probable assessment of deception dramatically increases.

However, while the correlational data and tests of significance indicate the importance of realism to the evaluation of truthfulness, the factor analyses of our best seven items suggest a rather different perspective. If one accepts at face value the data presented in Table 4, realism is by no means the most important variable in assessing a statement’s truthfulness or untruthfulness. Indeed, the factor analysis on the seven items in the truthful condition suggests that the four situational variables, clarity, temporal, spatial, and contextual detail, are all more important in the assessment of truthfulness than the three judgment variables, relevance, realism, and reconstructability (i.e., the four situational variables load more highly on the one extracted factor). These data suggest that the foremost hallmark of a true statement could be the mundane even trifling little situational details included in it. Such factual details which have little or no meaning in and of themselves, when included in a person’s statement may be emblematic of truth.

In addition, we find it equally interesting that in the factor analysis for the seven items in the untruthful condition, the
reconstructability variable has the single highest correlation with the untruthful factor. While realism remains relatively low in this factor analysis, reconstructability and realism still represent judgments about a statement that are conceptually related. Realism is fundamentally about the believability of a statement, while reconstructability refers to its plausibility and logical coherence. So in contrast to the assessment of truthful statements, our data suggest that the evaluation of untrue statements may have less to do with the situational attributes than with the judgment that the statement is just too inconsistent, too contradictory, and too implausible to be possible.

Rather than risking further over interpretation of our results, the most important point, in our opinion, is the fact that the 2 seven-item factor analyses both yielded single-factor solutions. As conceptually different as spatial information is from the judgment of the realism, all seven items in both conditions yielded single-factor solutions, indicating that these seven items are all measuring the same thing and are connected in some underlying fashion. Obviously, it is our belief that the theme underlying the factor analytic data is the truthfulness versus untruthfulness of the statements being evaluated. While we have yet to demonstrate the generalizability of our measure, its reliability and internal validity with regard to the present data seem clear. Our hope is to do further research on our seven best items to assess the extent to which they possess external as well as internal validity.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. Psychological Review, 88, 67-85.

Johnston, S. A., Powers-green, D., Brady, G., & Rahmani, S. (2013). Detecting deception in psychological evaluations. One-speaker paper presentation from 2013 Oregon Psychological Association Annual Conference, Eugene, OR. May 2013.

Roma, P., San Martini, P., Sabatello, U., Tatarelli, R., & Ferracuti, S. (2011). Validity of criteria-based content analysis (CBCA) at trial in free-narrative interviews. Child Abuse & Neglect, 35, 613-620.

ten Brinke, L., & Porter, S. (2013). Discovering deceit: Applying laboratory and field research in the search for truthful and deceptive behaviour. In B. Cooper, M. Ternes, & D. Griesel (Eds.), Applied issues in investigative interviewing, eyewitness memory, and credibility assessment. New York, NY: Springer. pp 221-237.

Trivers, R. (2011). The folly of fools: The logic of deceit and self-deception in human life. New York, NY: Basic Books.

Vrij, A. (2008). Detecting lies and deceit: Pitfalls and opportunities. Chichester, UK: John Wiley.

Vrij, A., Mann, S., & Fisher, R. P. (2006). An empirical test of the behaviour analysis interview. Law and Human Behavior, 30, 329-435.

Author Biographies

Dr. Shawn Adair Johnston is a forensic psychologist with approximately 40 years of experience, currently teaching at Portland State University.

Alexis Candelier graduated summa cum laude from Portland State University in December 2013 and is assisting Dr. Johnston in his current research program.

Dana Powers-Green and Syeda Rahmani are two post-baccalaureate students from Portland State University, who have been working with Dr. Johnston for the last two years on the research study.