The Dimensions of Deception Detection:

Self-reported Deception Cue Use is Underpinned by Two Broad Factors

(Word count: 6184)

Iain Hamlin (PhD, Lancaster University)

University of Strathclyde, Glasgow, UK, (p) +44 (0)141 553 6004, iain.hamlin@strath.ac.uk

Gordon R.T. Wright (PhD, University of London)

Goldsmiths, University of London, London, UK, (p) +44 (0) 207 078 5465, (f) +44 (0)20 7919 7873, g.wright@gold.ac.uk

Sophie Van der Zee (PhD, Lancaster University)

Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 300 DR Rotterdam, Netherlands

(p) 0031 104082776, s.c.vander.zee@ese.eur.nl

Stuart Wilson (PhD, University of Edinburgh)

Queen Margaret University, Edinburgh, UK, (p) +44 (0)131 474 0000, (f) +44 (0)131 474 0001 swilson@qmu.ac.uk

*Corresponding author. Address: Department of Strategy and Organisation, Strathclyde Business School, University of Strathclyde, 199 Cathedral Street, Glasgow, UK, G4 0QU

Email address: iain.hamlin@strath.ac.uk

Abstract

Many perceived cues to deception have been reported in the literature, but little attention has been paid to how they are combined when making deception judgments. The present research used a data-driven approach to investigate how cues are integrated when evaluating veracity. 215 participants performed a deception detection task before completing a deception cue use questionnaire. A factor analysis of the questionnaire data produced two dimensions, one associated with nonverbal cues to anxiety, the other associated with detail and level of involvement. The present work extends our understanding of deception detection processes and underlines the importance of examining not just the cues that people use but also the way in which they use them.

Keywords: deception detection; nonverbal cues; verbal cues

The Dimensions of Deception Detection:

Self-reported Deception Cue Use is Underpinned by Two Broad Factors

Many behavioral cues have been implicated in veracity judgments. Research on beliefs about cues to deception has revealed that people believe a wide variety of behaviors can be used to inform deception judgments, including nonverbal cues such as postural shifts, self-touching and lack of eye contact (Global Deception Research Team, 2006; Granhag, Anderson, Stromwall, & Hartwig, 2004; Zuckerman, Koestner, & Driver, 1981), paraverbal cues such as speech rate and voice pitch (Vrij & Semin, 1996) and linguistic cues such as plausibility (Vrij, Akehurst and Knight, 2006) and inconsistencies in content (Global Deception Research Team, 2006; Granhag et al., 2004; Vredeveldt, van Koppen, & Granhag, 2014). Furthermore, a recent meta-analysis of implicit deception cues produced a list of over forty behaviors which participants in deception studies had utilised in their deception judgments (Hartwig & Bond, 2011).

It would be cognitively prohibitive to evaluate each of these cues individually when making veracity judgments. It is likely that, as with other social judgments, the structure of deception cue use is organized along a smaller number of broad dimensions. For example, the structure of implicit personality judgments appears to be two dimensional in nature (Cooper & Hamlin, 2005; Rosenberg, Nelson, & Vivekananthan, 1968), as does the structure of the social perception of faces (Sutherland et al., 2013; Todorov, Said, Engell, & Oosterhof, 2008), the social perception of voices (McAleer, Todorov, & Belin, 2014) and the perception of emotion (Katsikitis, 1997; Takehara & Suzuki, 2001). Scant attention, however, has been paid to the dimensions underlying veracity judgments.

More direct evidence that deception cue use might be organised along a small number of broad dimensions comes from recent research by Hartwig and Bond (2011). Their meta-analysis of implicit deception cues suggested that deception judgments are more strongly related to global impressions of broad behavioral categories, such as 'plausibility' and 'verbal and vocal involvement', than the low-level behaviors that make-up these categories. Although veracity judgments must by necessity draw on low-level behaviors, the results of Hartwig and Bond suggest that these low-level behaviors are aggregated into broader dimensions when making judgments. However, the broad categories examined by Hartwig and Bond might only be proxies for the broad dimensions that people actually use. The number and composition of these broad dimensions remains an open question.

Research studies on deception cue use often cluster cues into broad categories, such as the extent to which participants use 'verbal' versus ‘nonverbal’ cues. For example, Mann, Vrij, and Bull (2004) found that police officers successful at detecting deception reported utilising verbal cues more frequently than less successful officers, whereas self-reported use of non-verbal cues was negatively related to deception detection accuracy. However, clustering cues in this way does not necessarily represent how cues are actually combined when making veracity judgments. For example, the extent to which an individual uses one nonverbal cue when judging deception is not necessarily related to the extent to which they use other nonverbal cues. Similarly, other ways in which deception cue use has been categorized, such as into the vocal, verbal and nonverbal domains (DePaulo, Lassiter & Stone, 1982; Feeley & Young, 2000) or vocal, verbal, facial and bodily domains (Porter, McCabe, Woodworth, & Peace, 2007), are again merely ad hoc categorisations and don't necessarily reflect even the researchers' conception of how veracity perceptions are structured, let alone how they are actually structured.

There are theoretical reasons to assume that deception cue use will be underpinned by verbal and nonverbal dimensions. The situational familiarity hypothesis (Stiff et al., 1989) posits that judges use verbal cues only in familiar situations – that is, situations in which they are in a position to evaluate the validity of verbal content. In unfamiliar situations, where it is more difficult to evaluate the validity of verbal content, deception judges rely more heavily on nonverbal cues. In support of the situational familiarity hypothesis, Reinhard, Sporer, Scharmach, and Marksteiner (2011) reported that the more familiar participants were with a deception context, the higher was their use of verbal content information. If verbal cues are more strongly utilized in deception contexts with high situation familiarity, and if deception judges vary in terms of their familiarity with the deception context, then verbal cues – and, by extension, nonverbal cues – should covary with each other in deception judgments and so potentially form verbal and nonverbal dimensions. Although a possible implication of the work of Stiff et al. and Reinhard et al. is that cue use covaries within modality, an empirical investigation is required to address how cues are actually combined in the process of deception judgment.

An understanding of how cues are combined when judging deception would be an important component of our understanding of why deception detection performance is generally so poor and, by extension, how it can be improved. The accuracy rate of participants in deception detection studies is typically only around 54%, where a 50% accuracy rate would be expected by chance (Bond & DePaulo, 2006). However, there is evidently scope to improve this. The fact that statistical models based on measurements of truthful and deceptive behaviors do significantly better than humans at predicting veracity suggests that it is possible for humans to make significant gains in deception detection performance (Hartwig & Bond, 2014; Vrij, Edwards, Robert, & Bull, 2000; Vrij, Akehurst,

Soukara, & Bull, 2004). Moreover, a recent meta-analysis demonstrated that deception detection performance can be enhanced through training (Hauch, Sporer, Michael, & Meissner, 2014), and the effect of training on performance has been found to be attributable to more than just the 'placebo' effect of receiving training (Levine, Feeley, McCornack, Harms, & Hughes, 2005). There is considerable scope to improve performance, and attempts at altering deception detection processes would benefit from an understanding of how these processes work, including how cues are integrated in decision making.

Extending previous research on deception cue use, the present research uses a data-driven approach to investigate how deception cues are combined during the process of deception detection. This approach has proven useful in other areas of social perception research. For example, using a data-driven approach, Oosterhof and Todorov (2008) derived a structural representation of explicit personality perceptions by asking participants to rate faces on several personality traits and then subjecting these ratings to a principal components analysis. The resulting dimensions, which mapped on to the well-established two factor model of trustworthiness and dominance, were interpreted by Oosterhof and Todorov as the main dimensions underpinning personality judgments. Similarly, in the domain of psychoacoustics, the structure underlying voice perception has been examined by applying multidimensional scaling analysis to similarity ratings of pairs of voices (Murry & Singh, 1980; Shrivastav, 2006; Singh & Murry, 1978). The present research adopted a factor-analytic approach, similar to that used by Oosterhof and Todorov, to recover the structure of deception cue use underpinning veracity judgments.

Although this study is largely exploratory, we hypothesise that the structure of deception cue use will be dimensional in nature. That is, we anticipate that participants will not evaluate individual cues independently from each other, but will systematically integrate

cues when making judgments. That deception judges combine behavioral cues has been implicitly suggested by previous research which has found larger correlations between broad cues to deception and veracity judgments than between narrow cues to deception and veracity judgments (Hartwig & Bond, 2011). Furthermore, we hypothesise that, as in other areas of social perception research, cue use will be underpinned by a very small number of broad factors. We also hypothesise that, based on dual process theories of veracity judgments, nonverbal cues and verbal cues will be associated with different factors.

Method

Participants

220 university students (69 males) participated in a single experimental session for a small monetary reward. They were recruited over Prolific Academic, an online participant recruitment platform. The mean age of participants was 39.25 years (SD = 13.09). 5 participants were dropped for failing an attention check embedded in the cue-use questionnaire, leaving a sample of 215 participants.

Video stimuli. 26 videos were used as stimuli in the current study. Each of the videos, which were between 8 and 20 seconds long, featured a student from a British university either lying or telling the truth in response to one of four questions: ‘Tell me about a holiday you went on in the past’, ‘Tell me about a plan you have for the future’, ‘Tell me about an interest you have in life’, ‘Tell me about something you don’t like or are frightened of’. 13 unique individuals were featured in the videos, each one telling one lie and one truth. Consequently, half of the videos showed participants lying, half depicted participants telling the truth.

Deception cue use questionnaire. A deception cue use questionnaire was created to measure the extent to which participants believed they utilised specific behaviors when making their veracity judgments. The cues used in the questionnaire were those reported in Hartwig and Bond's (2011) meta-analysis as having a statistically significant relation to perceived deception. Each questionnaire item consisted of the label for the cue given by Hartwig and Bond and the attendant description of the cue given in Appendix A of DePaulo et al. (2003). The cues 'Nonverbal deception pose' and 'Verbal deception pose' were not included in the questionnaire because no further description was offered of them by Hartwig and Bond, and their meaning was not self-evident. The cue 'Total disturbances (ah and non-ah speech disturbances)' was also not included in the questionnaire, because, in Hartwig and Bond's meta-analysis, 'Non-ah disturbances' did not have a statistically significant relation to perceived deception, and 'Ah disturbances' was already included as an item in the questionnaire. A total of 39 items were thus included in the questionnaire.

Participants were asked to rate how important each of the behaviors was to them when making their judgments. Ratings were made on a 5-point scale (*not at all important* to *very important*). After each rating, participants indicated whether they associated increases in the behavioral cue with deception or truthfulness. For example, after rating how important 'response length' was to them when making their judgment, they responded to the statement 'the longer their response length, the more I thought they were…' by selecting '*lying'*, '*telling the truth'*, or '*neither: I answered '1' for the above question*'.

Procedure

Participants completed a short deception detection task created using the video stimuli described above. The computer-based task, run using Qualtrics, displayed a video clip

followed by an on-screen request to rate whether the person in the video was lying or telling the truth from 1 (*Definitely lying*) to 6 (*Definitely telling the truth*). After answering the question, participants clicked 'Next' to view the next video. This process continued until all 26 videos had been viewed and rated. The order in which videos were displayed was randomised across participants. Following completion of the task, participants completed the deception cue use questionnaire.

Deception judges had an overall deception detection accuracy rate of 52.8% (SD = 9.6%), which is comparable to the accuracy rate generally reported in the literature (Bond & DePaulo, 2006). In line with the results of the majority of deception detection studies, participants were more successful at detecting truths (57.7%, SD = 16.7%) than lies (48%, SD = 17.3%). Participants displayed a slight truth bias, judging 54.9% of messages as true.

Results

It is possible that the self-reports of cue use gathered at the end of the deception detection sessions did not accurately reflect the extent to which the cues were utilised during the session. To check that online, judgment-by-judgment reports of cue use are indeed strongly related to final, aggregate self-reports, we recruited a separate sample of 34 participants from the same population and asked them to make both judgment-by-judgment ratings of cue use for 13 deception videos and aggregate self-reports at the end of the session. We gave participants only 13 of the 26 videos (13 unique individuals) because if participants rated their cue use for all 26 videos, then they would had to have made a prohibitively large number of cue use ratings (945).

The mean correlation between the averaged judgment-by-judgment ratings and the aggregate self-reports provided at the end of the session was large (mean r = .51; SD = .18),

suggesting that end-of-session, aggregate self-reports were accurate reflections of how participants would report cue use on a judgment-by-judgment basis.

Structure of deception cue use

In order to uncover the broad behavioral dimensions that underpinned participants' veracity judgments, an unweighted least squares factor analysis was performed on the deception cue use questionnaire data. Factor analysis is a statistical procedure which is generally considered to require large sample sizes in order to produce reliable results. However, the issue of absolute sample size in factor analysis is an often misunderstood one, with rules of thumb for necessary sample sizes varying widely and apparently often based on little other than the personal experience of the proponent (MacCallum, Widaman, Preacher, & Hong, 2001; MacCallum, Widaman, Zhang, & Hong, 1999). The few empirical studies that have investigated the issue of sample size on factor recovery have concluded that rules of thumb are not valid or useful: there is no absolute sample size or ratio of items to participants that is sufficient to recover a set of population factors (de Winter, Dodou, & Wieringa, 2009; MacCallum et al., 1999; 2001; Mundfrom, Shaw, Lu Ke, 2005).

In general, the most important factor influencing the accurate recovery of factors appears to be the ratio of variables to factors. If each factor has at least 6 or 7 items loading highly on it, then the factor solution is almost always robust, even with modest sample sizes in the region of 100-200 (MacCallum et al., 2001). Even with low communalities, if the ratio of variables to factors is high, then population factors are almost always accurately recovered (MacCallum et al., 1999; Mundfrom et al., 2005). The structure of social judgments tends to be characterised by a very small number of broad dimensions. We therefore anticipated that the items of the questionnaire would be split between a small number of consequently

overdetermined factors, resulting in a factor structure robust enough to be recovered with a sample size of 215.

Because the cues rated by participants during the study were not selected on the basis that they had previously been found to be relevant to veracity judgments of those particular videos, it was possible that one or more of the cues were generally not used by participants when judging the videos. If some of the cues were not utilised by participants, then the distribution of ratings of such cues would be heavily skewed, leading to an artefactual factor solution. Therefore, the kurtosis value of each cue was examined. Four cues displayed high kurtosis because they were rated as being very rarely used: 'Pleasant face', 'Pleasant voice', 'Attractive', and 'Babyface'. These four cues all had kurtosis values of above .9, much higher than the kurtosis values of any other cues (the next three highest values were .409, .319, and .314). Consequently, ratings of these four cues were not entered into the factor analysis.

It is possible that different participants might both rate the same cue highly, but use the cue in different ways. That is, some participants might rate a cue highly because they think it is a cue to deception, whereas as others might rate it highly because they think it is a cue to truthfulness. In order to take directionality of cue use into account, some simple transformations were performed on the data. If a participant stated that they associated an increase in a cue with deception, then 4 was added to their cue use rating. If a participant stated that they associated an increase in a cue with truthfulness, then their cue use rating was subtracted from 6. If the participant indicated that they associated neither deception nor truthfulness with the cue, then their cue use rating became '5'. The result of these transformation was that participants' cue use ratings were now directional, ranging from 1 to 9 – that is, from strongly associating the cue with truthfulness to strongly associating the cue with deception.

In line with our expectation that the structure of deception cue use would be dimensional, Bartlett's test of sphericity was significant ($\chi 2(595) = 2027$, $p < .001$) and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .79, suggesting that the data were dimensional in nature and thus suitable for factor analysis (values of 0.5 or above are generally held to indicate suitability for factor analysis; Hair, Anderson, Tatham, & Black, 1995; Tabachnick & Fidell, 2001). We used various statistical methods to decide on the appropriate number of factors to extract. An examination of the eigenvalues revealed that the initial two eigenvalues were significantly larger than all proceeding eigenvalues (the first six eigenvalues of the reduced correlation matrix were 7.63, 2.14, 1.19, 0.89, 0.77 and 0.72 respectively). Similarly, a parallel analysis of eigenvalues suggested that a two-factor solution was most appropriate, as did Schwarz's Bayesian Information Criterion (BIC) dimensionality test. Taking the results of these statistical tests together, we decided that a two-factor solution best represented the data. The presence of two broad factors was in line with our expectation that, similar to other types of social perception, the perception of deception would be underpinned by a very small number of broad factors. Combined, the two factors accounted for 31.8% of the variance in the dataset. Because there was no theoretical reason to assume that the factors would be orthogonal, promax was chosen as the method of rotation.

[Table 1 about here]

Table 1 presents factor loadings on the two cue use questionnaire factors. Factor 1, *Difficulty*, contained items related to nonverbal and paraverbal displays of anxiety (e.g. 'Nervous', 'Gaze aversion', 'Postural shifts', and 'Ah disturbances'). Factor 2, *Involvement*, contained items related to involvement and detail (e.g., 'Verbal and vocal involvement', 'Details', 'Expressive face' and 'Hand gestures'). The two factors were highly correlated (r =

0.63), suggesting that the extent to which participants systematically relied on nonverbal cues in their veracity judgments was strongly related to the extent that they systematically relied on verbal cues.

Relationship between cue use and detection accuracy

To investigate whether participants' cue use influenced their deception detection performance, factor scores on the two cue use dimensions and self-reported use of each individual cue were entered into correlation analyses with detection accuracy (dprime) and truth bias scores. Table 2 presents the correlations between cue use and both detection accuracy and truth bias. The extent to which participants drew on the two cue use dimensions was unrelated to both accuracy at detecting deception and the amount of truth bias they exhibited.

[Table 2 about here]

Of the individual cues, only 'response latency' and 'vocal uncertainty' were significantly correlated with detection accuracy, suggesting that the self-reported use of specific cues was at best only very weakly related to detection accuracy. It should be noted that both correlations are negatively valenced. That is, the more that participants reported drawing on response latency and vocal uncertainty when judging veracity, the less accurate their judgments were. However, caution is urged when interpreting these significant results, owing to the large number of correlations examined and the consequent possibility of type 1 errors. No cues were associated with truth bias.

Discussion

The present study sought to uncover how people integrate behavioral cues when making veracity judgments. The results of previous research had suggested that low-level behaviors are combined into broad behavioral categories when evaluating deception (Hartwig & Bond, 2011). However, it was unclear whether these reported categories were simply proxies for the actual behavioral dimensions used by people to inform their veracity judgments. We found that two very broad dimensions underpinned the self-reported use of behavioral cues in deception judgments. One dimension was composed of items related to nonverbal and paraverbal signs of anxiety, the other was composed of items related to verbal involvement and amount of detail. The dimensions were very broad: each contained over a dozen different types of behavior. The breadth and content of these dimensions suggests that each represents a global, impressionistic evaluation of behavior, one a global evaluation of nonverbal and paraverbal cues to anxiety, the other a global evaluation of detail and involvement, with each dimension underpinned by many specific, low-level behaviors.

These findings are in line with the results of Hartwig and Bond's (2011) meta-analysis of implicit cues to deception. Similar to their findings, our results suggest that deception judgments are underpinned by broad impressions of behavior. Hartwig and Bond reported that deception judgments are informed by broad impressions of behavioral categories such as 'plausibility' and 'details'. The present research indicates that these categories are lower-level proxies of the significantly broader behavioral dimensions that people use when judging deception.

The issue was previously raised as to whether the structure of deception cue use maps onto any of the categorisations of cues typically used by deception researchers, such as 'verbal and nonverbal' or 'verbal, vocal, bodily and facial'. Interestingly, the structure of self-reported cue use appears to at least partly fit into the verbal-nonverbal system of

categorisation used by some researchers to structure their theories, methodologies, or data analysis. The *Involvement* factor was largely verbal in nature and the *Difficulty* factor was largely nonverbal, providing some evidence for the validity of the verbal-nonverbal distinction and suggesting that the main ways in which deception judges vary from each other is in the importance they place on nonverbal and verbal cues. However, it should be noted that it would be overly simplistic to conceptualise the two factors purely in terms of modality. Some of the items loading highly on the 'verbal' dimension (e.g. ‘Speech rate’ and ‘Hand gestures’) are paraverbal or nonverbal in nature. Moreover, the lack of verbal items loading on the 'nonverbal' dimension is perhaps simply an artefact of the relative lack of perceived verbal cues to anxiety.

Neither of the two dimensions were related to deception detection accuracy. This was perhaps a reflection of the weak relationships between the use of specific deception cues and detection accuracy. Almost none of the specific self-reported cues were significantly correlated with detection accuracy, so it is not surprising that global judgments based on these cues were similarly unrelated to accuracy. One of the probable reasons for the lack of significant relationships between cue use and detection accuracy was the overall low detection rate in the present study: participants performed at around chance levels when judging veracity. A context in which detection rates are higher than in the present study should provide more fertile ground for investigating the relationships between cue use and deception detection performance.

The present study may offer some insight into the reasons for the poor deception detection performance of humans widely reported in the literature (Bond & DePaulo, 2006; Kraut, 1980; Vrij, 2008). It is interesting to note that, although several aspects of deception phenomenology appear to find their counterpart in the deception detection dimensions

reported in this research (e.g. the anxiety and cognitive load factors described by Zuckerman, DePaulo & Rosenthal (1981)), observers appear not to pay much attention to the possibility that targets might adopt behavior control strategies. That is, even though liars often implement one or more deception strategies when perpetrating their lies (Colwell, Hiscock-Anisman, Memon, Woods, & Milchlik, 2006; Hartwig, Granhag, & Stromwall, 2007; Stromwall, Hartwig, & Granhag, 2006; Stromwall & Willen, 2011), deception judges do not appear to look for cues to these strategies in targets whose veracity they are evaluating.

Perhaps this failure by deception judges to look for cues to deception strategies occurs because the pathways between deception strategies and cues to deception are not as clear in judges' minds as the arguably more straightforward pathways between anxiety, cognitive load and cues to deception. Another possibility is that the potential behavioral effects of strategies are not paid attention to by deception judges because they most likely assume that these strategies, if employed, will result in behavior that is indistinguishable from truths. Deception strategies are a causal factor in the production of cues to deception (Vrij, Semin, & Bull, 1996), so by failing to look for signs of deception strategies, participants ignore many potentially useful cues to deception. This would be expected to have a detrimental effect on their deception detection performance. Consequently, deception detection training efforts might benefit from complementing current patterns of cue utilisation by directing trainees toward cues to the manifestation of deception strategies.

Conversely, an alternative approach to improving the accuracy of veracity judgments is to align training procedures with judges' natural way of processing deception cues. So, because deception detectors appear to base their judgments on global evaluations of difficulty and involvement, training effort should perhaps be focussed along these dimensions. Future

work would be well-directed to examine which of these two types of training is most efficacious.

Several limitations of the present research are worth noting. First, it remains unclear whether the two factors reported will generalise across different types of lie. However, the two factors are very broad in terms of content and are consequently of relevance to a great many deceptive situations, so it is expected that they will demonstrate a high degree of replicability. Future research should establish how stable the two factors reported here are across different deception contexts and investigate the extent to which their adoption promotes deception detection accuracy.

Second, the present study investigated the structure of self-reported deception cue use rather than implicit cue use. It is possible that there is a dissociation between the cues that people report using to inform their veracity judgments and the cues that they actually (i.e. implicitly) use (Hartwig & Bond, 2011). However, there is little evidence in the deception literature to suggest that such a dissociation exists. Indeed, the extant literature suggests that the opposite is true. For example, Zuckerman, Koestner, et al. (1981) reported a moderate correlation between the effect sizes for implicit cues to deception and the corresponding effect sizes for self-reported cues to deception, suggesting that self-reported cues to deception map well onto the cues that people actually use when making deception judgments. Moreover, self-reported cue use has been reported to predict the accuracy of veracity judgments (Reinhard, Dickhäuser, Marksteiner, Sporer, 2011). Similarly, in a study of deception detection among trainee teachers, each of the sixteen self-reported cues to deception measured were found to be significant predictors of veracity judgments (Marksteiner et al., 2012). There is substantial evidence that self-reported cue use accurately reflects implicit cue use.

Indeed, the findings of the present research lend further credence to the idea that there is no great disparity between the self-reported and actual behaviors used in veracity judgments. In support of a disparity, Hartwig & Bond (2011) cite the fact that the implicit cues that most strongly correlate with veracity judgments are broad and impressionistic, whereas the narrower cues commonly featured in self-reports of cue use, such as leg movements and postural shifts, correlate relatively poorly with veracity judgments. However, this pattern of results should be expected if broad dimensions underpin veracity judgments. Our results suggest that veracity judgments are based on broad, impressionistic dimensions, which are made-up of narrower cues, such as hand gestures, gaze aversion and response length. As individual, constituent elements of the broad dimensions, the narrow cues should be expected to have weaker relationships with veracity judgments than the broad dimensions do. To say that the narrow behavioral cues typical of self-reports are not strongly implicated in veracity judgments would be to misunderstand the mediated nature of their relationship with veracity judgments.

Finally, two psychometric issues merit discussion. The two factors extracted in the present study together accounted for 31.8% of the variance in scores on the questionnaire. This value is lower than in most other factor analyses. For example, in a meta-analysis of the amount of variance accounted for in factor analyses, Peterson (2000) reported that factor analyses of over 31 items accounted for on average 48.1% of the variance in the data. There are several reasons why the factor analysis reported in the current study accounted for less variance than the average factor analysis. Unlike in most factor analyses, our data were not pre-structured to increase the values of the loadings on the resultant factors, and, by extension, increase the amount of variance accounted for by the factor solution. Additionally, the percentage of variance explained by a factor solution is inversely related to the number of

items entered into a factor analysis (Peterson, 2000), so the relatively large number of items used in our analysis negatively impacted the percentage of variance accounted for. It should also be noted that the relatively small percentage of variance accounted for might also in part be explained by psychological factors. In particular, it most likely reflects idiosyncrasy in participants' deception cue use beyond the two broad cue use dimensions extracted in the factor analysis.

A second psychometric issue pertains to the significant number of items with relatively low factor loadings. Although this would be a serious issue when performing factor analysis in aid of scale construction, it is not inherently problematic in the context of the present research. Indeed, it is to be expected that some of the behavioral cues studied will be only weakly drawn on by global judgments of anxiety and involvement, and so, as a result, will have low loadings on the two factors.

In conclusion, the present study extended previous research on deception detection by using a data-driven, factor-analytic approach to describe how behavioral cues are combined when judging deception. Two broad factors underpinned veracity judgments, one related to nonverbal and paraverbal cues to anxiety, the other defined by involvement and detail. The results extend our knowledge of deception detection processes and underline the importance of examining not just the cues that people use when making veracity judgments but also the manner in which they use them.

References

Bond, C. F. & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234.

Colwell, K., Hiscock-Anisman, C., Memon, A., Woods, D., & Michlik, P. (2006). Strategies of impression management among deceivers and truth-tellers: How liars attempt to convince. *American Journal of Forensic Psychology*, *24*, 31-38.

Cooper, C. & Hamlin, I. (2005). Observers spontaneously use intelligence, extraversion, neuroticism and psychoticism when evaluating personality. *Personality and Individual Differences, 39*, 123-130.

DePaulo, B. M., Lassiter, G. D., & Stone, J. I. (1982). Attentional determinants of success at detecting deception and truth. *Personality and Social Psychology Bulletin*, *8*, 273-279.

DePaulo, B. M., Lindsay, J. L., Malone, B. E., Muhlenbruck, L., Charton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulleting, 129,* 74-118.

De Winter, J. C. F, Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research, 44*, 147-181.

Feeley, T. H., & Young, M. J. (2000). The effects of cognitive capacity on beliefs about deceptive communication. *Communication Quarterly, 48*, 101-119.

Global Deception Research Team (2006). A world of lies. *Journal of Cross-Cultural Psychology, 37*, 60-74.

Granhag, P. A., Andersson, L. O., Stromwall, L. A., & Hartwig, M. (2004). Imprisoned knowledge: Criminals' beliefs about deception. *Legal and Criminological Psychology, 9*, 103-119.

Hair, J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis*. 4th ed. New Jersey. Prentice-Hall Inc.

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*, 643–659.

Hartwig, M., & Bond, C. F. (2014). Lie Detection from Multiple Cues: A Meta-analysis. *Applied Cognitive Psychology, 28(5)*, 661-676.

Hartwig, M., Granhag, P.A., & Strömwall, L.A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, *13*, 213-227.

Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2014). Does training improve the detection of deception? A meta-analysis. *Communication Research*, 1-61.

Katsikitis, M. (1997). The classification of facial expressions of emotion: A multidimensional-scaling approach. *Perception, 26,* 613–626.

Kraut, R. E. (1980). Humans as lie detectors: Some second thoughts. *Journal of Communication, 30*, 209-216.

Levine, T. R., Feeley, T., McCornack, S. A., Harms, C., & Hughes, M. (2005). Testing the effects of nonverbal training on deception detection accuracy with the inclusion of a bogus training control group. *Western Journal of Communication, 69,* 203-218.

MacCallum, R. C., Widaman, K. F., Preacher, K. L., & Hong., S. (2001). Sample size and factor analysis: The role of model error. *Multivariate Behavioral Research, 36,* 611-637.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4,* 84-99.

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

Marksteiner, T., Reinhard, M. A., Dickhäuser, O., & Sporer, S. L. (2012). How do teachers perceive cheating students? Beliefs about cues to deception and detection accuracy in the educational field. *European Journal of Psychology of Education, 27,* 329-350.

McAleer P., Todorov A., & Belin, P. (2014). "How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices." *PLoS ONE 9(3):* e90779. DOI: 10.1371/journal.pone.0090779

Mundfrom, D., Shaw, D., & Ke, T. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing, 5(2)*, 159-168.

Murry, T., & Singh, S. (1980). Multidimensional analysis of male and female voices. *The Journal of the Acoustical Society of America, 68*, 1294–1300.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the USA, 105*, 11087-11092.

Peterson, R. A. (2000). A Meta-Analysis of Variance Accounted for and Factor Loadings in Exploratory Factor Analysis. *Marketing Letters, 11*, 261-275.

Porter, S., McCabe, S., Woodworth, M., & Peace, K. A. (2007). Genius is 1% inspiration and 99% perspiration...or is it? An investigation of the effects of motivation and feedback on deception detection. *Legal and Criminological Psychology, 12*, 297–309.

Reinhard, M. A., Dickhäuser, O. Marksteiner, T., & Sporer, S. L. (2011). The case of Pinocchio: Teachers ability to detect lies and truth. *Social Psychology of Education, 14*, 299-318.

Reinhard, M. A., Sporer, S. L., Scharmach, M., & Marksteiner, T. (2011). Listening, not watching: Situational familiarity and the ability to detect deception. *Journal of Personality and Social Psychology, 101(3)*, 467-484.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology, 9*, 283–294.

Shrivastav, R. (2006). Multidimensional Scaling of Breathy Voice Quality: Individual Differences in Perception. *Journal of Voice, 20*, 211–222.

Singh, S., & Murry, T. (1978). Multidimensional classification of normal voice qualities. *The Journal of the Acoustical Society of America, 64*, 81–87.

Stiff, J. B., Miller, G. R., Sleight, C. Mongeau, P., Garlick, R., & Rogan, R. (1989). Explanations for visual cue primacy in judgments of honesty and deceit. *Journal of Personality and Social Psychology, 56*, 555-564.

Strömwall, L. A., Hartwig, M., & Granhag, P. A. (2006). To act truthfully: Nonverbal behavior and strategies during a police interrogation. *Psychology, Crime & Law*, *12*, 207-219.

Strömwall, L.A., & Willén, R.M. (2011). Inside criminal minds: Offenders' strategies when lying. *Journal of Investigative Psychology and Offender Profiling, 8*, 271-281.

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J. M., Burt, M. D. & Young, A W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition, 127*, 105-118.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics.* Boston: Allyn and Bacon.

Takehara, T. & Suzuki, N. (2001). Differential processes of emotion space over time. *North American Journal of Psychology, 3*, 217–228.

Todorov, A., Said, C. P., Engell, A. D. and Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in cognitive sciences, 12*, 455–460.

Vredeveldt, A., Van Koppen, P. J., & Granhag, P. A. (2014). The inconsistent suspect: A systematic review of consistency in truth tellers and liars. In R. H. Bull (Ed.), *Investigative interviewing* (pp. 183–207). London: Springer.

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

Vrij, A., Akehurst, L. & Knight, S. (2006). Police officers', social workers', teachers' and the general publics' beliefs about deception in children, adolescents and adults. *Legal and Criminological Psychology, 11*, 297-312.

Vrij, A., Akehurst, L, Soukara, S. & Bull, R. (2004). Detecting Deceit Via Analyses of Verbal and Nonverbal Behavior in Adults and Children,. *Human Communication Research, 30,* 8-41.

Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting Deceit Via Analysis of Verbal and Nonverbal Behavior. *Journal of Nonverbal Behavior, 24*, 239–263.

Vrij, A. & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior, 20*, 65-80.

Vrij, A., Semin, G. R., & Bull, R. (1996). Insight into behavior during deception. *Human Communication Research, 22*, 544–562.

Zuckerman, M., DePaulo, B.M., & Rosenthal, R. (1981). Verbal and non-verbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology, volume 14* (pp. 1-57). New York: Academic Press.

Zuckerman, M., Koestner, R., & Driver, R. (1981). Beliefs about cues associated with deception. *Journal of Nonverbal Behavior, 6*, 105-114.

Table 1

Factor Loadings of the Questionnaire Items on Two Factors

	Questionnaire factor	
Questionnaire item	Difficulty	Involvement
Postural Shifts	.75	.34
Shrugging	.73	.35
Gaze aversion	.72	
Weird behaviors	.66	
Nervous	.65	
Eye contact	-.61	
Blinking	.61	
Ah disturbances	.58	
Vocal uncertainty	.5	
Not spontaneous	.37	
Nonfluent	.37	
Verbal and vocal involvement		.75
Details		.74
Involved		.7
Expressive face		.55
Cooperative		.51
Arm movements	.41	.49
Response length		.48
Logical structure		.47
Hand gestures		.46
Response latency		-.46
Friendly		.43
Plausibility		.42
Speech rate		.42
Self-references		.4
Vocal immediacy		.38
Competent		.38
Realistic		.37
Unfilled pauses		-.37
Block access		.35
Indifferent		-.34

Note: Factor loadings below .32 are not displayed.

Table 2

Correlations Between Cue Use, Detection Accuracy (‘d) and Truth Bias (C)

Cue type	Questionnaire item	Accuracy (‘d)	Truth bias (C)
Cue use dimension	Dimension 1 – Difficulty	-0.06	-0.07
	Dimension 2 – Involvement	-0.06	-0.07
Verbal	Details	0.00	0.09
	Block access to information	-0.04	0.01
	Plausibility	-0.06	0.13
	Logical structure	-0.10	0.03
	Ambivalent	0.00	-0.08
	Verbal and vocal involvement	-0.03	0.08
	Self-references	-0.03	0.13
	Other references	0.02	0.02
	Realistic	-0.03	0.12
	Ingratiating	-0.06	0.01
Paraverbal	Response length	0.06	-0.02
	Response latency	-0.16*	-0.01
	Speech rate	-0.07	-0.04
	Involved	-0.04	0.03
	Vocal immediacy	-0.05	0.09
	Vocal uncertainty	-0.15*	0.06
	Unfilled pauses	-0.05	0.09
	Ah disturbances	-0.04	-0.03
	Nonfluent	-0.01	0.00
	Friendly	-0.10	0.03
	Cooperative	-0.05	0.08
	Indifferent	0.02	-0.01
	Not spontaneous	-0.06	0.03
	Competent	-0.13	0.03
Nonverbal	Expressive face	-0.10	0.09
	Eye contact	-0.05	-0.08
	Thinking hard	-0.05	-0.07
	Gaze aversion	-0.12	0.04
	Shrugging	-0.07	-0.02
	Blinking	0.02	-0.09
	Postural shifts	0.12	-0.05
	Hand gestures	-0.05	-0.09
	Arm movements	0.01	-0.09
	Nervous	-0.08	-0.13

Note: * denotes significance level above .05 (2-tailed).