Person Perception Aspects of Judgments of Truthfulness in Public Appeals

David Canter\textsuperscript{1}, Maria Ioannou\textsuperscript{2}, Donna Youngs\textsuperscript{3} and Gemma Chungh\textsuperscript{4}

\textsuperscript{1234}International Research Centre for Investigative Psychology (IRCIP), University of Huddersfield, Huddersfield, UK
Person Perception Aspects of Judgments of Truthfulness in Public Appeals

Abstract

Although detection of deception accuracy rates are well established, the person perception components that are the basis for these judgments remain unclear. To explore this, 30 academics’ person perceptions, as well as truthfulness judgement, of the individual presenting a televised appeal were measured using a 14 item scale. Twelve appeals (6 genuine and 6 false) for information regarding the whereabouts of a missing relative, or for information to apprehend the person who murdered their relative were used. The person perception scale consisted of 1) global, abstract judgments (open, deceptive, genuine, trustworthy and emotional) and 2) behavioural indices (facial pleasantness, facial animation, arousal, tension, involvement, verbal; consistency, plausibility and directness and vocal certainty). Multiple regression identified person perceptions of openness, (non)deceptiveness, genuineness, trustworthiness and verbal plausibility as significant predictors of truthfulness judgments. Future research should now explore the relation of these person perception components of truth judgments to the accuracy.

Keywords: Person perception, high stakes-lies, lie detection, judgments of truthfulness, false and genuine appeals, deception
Although Vrij (2008) identified that no less than 150 peer reviewed journal articles on lying behaviour or lie detection have been published annually during the last decade from disciplines including communication studies, psychology and criminal justice studies (c.f. Buller & Burgoon, 1996; Vrij, 2008; King & Dunn, 2010), much less consideration has been given to the qualities within an account that make it appear truthful. An unexplored issue here is person perception; the attributes that people perceive in the speaker that make them appear truthful and give their account veracity.

**Person perception**

In everyday life people form impressions and make judgments about each other, their personalities or form hypotheses about the kinds of person they are. According to social psychological theories, our behaviour towards others is based on our perceptions of their motives, temporary states and stable characteristics (Kraut & Poe, 1980). The study of person perception refers to the different mental processes that we use to form impressions of other people. This includes not just how we form these impressions, but the different conclusions we make about other people based upon our impressions.

Researchers have recognised that person perception is a constructive process in which expectations, predispositions and implicit personality theories are sometimes more important than the actual characteristics of the people we judge (Schneider, 1973). Social judgments, in general, and person-perception judgments, in particular, involve the imposition of our internal cognitive constructs on the complex and often indeterminate characteristics of others (Forgas & Bower, 1988; Heider, 1958; Kelly, 1955).
People often form impressions of others very quickly with only minimal information. A number of factors can influence these impressions such as roles and social norms, physical cues, verbal descriptions, salience, context effects, social categorization, implicit personality theories, behaviour and so on. We seem to make inferences about personality traits spontaneously and without reflection or deliberation (Hassin & Tropé, 2000; Todorov et al, 2005). Perceptual cues may act as triggers for categorisation of individuals, which may in turn lead to the activation of stereotypes (Cloutier, Mason & Macrae, 2005). In many cases the cues to category membership (sex or age, for example) are themselves accurate, even if the stereotypes they activated are not (Penton-Voak et al, 2006).

Of the many person attributes that influence our behaviour towards others, a number of theorists have argued that perceptions of sincerity and truthfulness are fundamental (Kraut & Poe, 1980). For example, in the case of ingratiations, Jones and Wortman (1973) argue that tactics such as self-presentation, flattery or opinion similarity will fail unless the target of the ingratiations attempt seems to them sincere. It is intriguing to consider the possibility that there may be perceptual cues that accurately inform judgments of truthfulness. In their study Todorov, Mandisodza, Goren and Hall (2005) found that inferences of competence made from politicians' faces predicted US congressional election outcomes at levels far above chance.

Researchers have found that facial appearance influences social judgments. We input trait qualities to people on the basis of their physical features, i.e. we expect baby-faced individuals to be warm, affectionate, honest and weak (Andreotti et al, 2001) and facially attractive individuals to be intelligent (Zebrowitz et al. 2002). There is also evidence that facial babyishness and age affect perceivers' impressions of the stimulus person's veracity.
Researchers examined whether these variables also influenced the credibility attributed to written statements purportedly made by these people in addition to several topics of interest in deception-detection research. When the statements truthful or deceptive were accompanied by baby-faced picture participants tended to judge them as truthful (Masip, Garrido, Herrero, 2003).

In Kraut and Poe's (1980) study of person perception, professional inspectors in the US customs service and laymen judged whether they wanted to search airline passengers going through a mock customs inspection. Results showed that judges' decisions to search travelers were based primarily on the travelers' comportment. Travelers were most likely to be searched if they were young and lower class, appeared nervous, hesitated before answering, gave short answers, avoided eye contact with the interviewer, shifted their posture and had taken pleasure trips. Individual differences among judges had little effect on the cues used. The results demonstrate the value of a social psychophysical approach to person perception that focuses on the behaviour of the perceived (Kraut & Poe, 1980).

In terms of veracity assessment, the person perception literature opens up questions then about the basis for people's judgments of truthfulness that are distinct from questions about the accuracy of these judgments. To date, this role of person perception factors in how truthfulness judgments are formulated is understudied. Some characteristics of the perceived, such as physical beauty (Berscheid & Walster, 1974), or nervousness may be powerful predictors of judgments across situations and judgment tasks. Identifying what these components are that contribute to a truthfulness judgment, and subsequently which predict its accuracy, could have important practical significance for law enforcement and other professionals' decision making (Kraut & Poe, 1980).
The Present Study

The present study explores two classes of person perception attributes associated with judgments of truthfulness in false and genuine public appeals for help with missing or murdered relatives: Behavioural (perceptions of the person's behaviour during the appeal) and Global (general characteristics of the person) attributes.

The Stimulus/Public Appeals

Invariably lie detection research utilises laboratory experimental designs. While this is justifiable, permitting the pursuit of analysing cause and effect, as deceit is intrinsically associated with human contact and interaction, its ecologically validity is highly questionable. Furthermore laboratory studies are not conducive to facilitating high risk social interaction invariably intrinsic to high stake deceptive behaviour (Koper & Sahlman, 1991).

Despite innovative designs to simulate real life events, often lie detection literature requires the induction of participants’ deceptive performance. To illustrate participants have been required to lie about the course they are studying at University (Vrij & Holland, 1998), have been induced to cheat and then to lie about it (deTurck & Miller, 1985; Feeley, 1996), provided with the opportunity of taking money and, then to lie about their actions in a later interview (Frank & Ekman, 1997) or the film they have just viewed (Ekman, Friesen & Scherer, 1976). However these designs carried very little at stake if the lie did not succeed.

Koper & Sahlman (1991) find such research problematic because of its inability to capture the personal involvement of the person required to lie in the experiment, which considerably reduces their motivation to succeed. Koper & Sahlman posit that this potentially undermines the entire premise of exploring deceptive behaviour in an experimental setting. However in
defence of the majority of existing lie detection research it has not sought to examine lies in high stake situations such as lying to police officers during an investigation; it has approached the issue from the position of everyday lies, although the findings are generalised to other situations above and beyond the initial scope of the research.

In many studies the deceptive behaviour is acted out, therefore the deception does not occur naturally. The present study is trying to overcome such problems by using public appeals. To the authors’ knowledge only two other studies used press conferences. In Vrij & Mann's (2001b) study police officers were exposed to five high-stake lies told by five different people. Videoclips were presented of press conferences given by five people whose relatives were missing or had been killed. All five people in this case were lying during these press conferences as each of them was subsequently found guilty of killing their own relatives. Wright Whelan, Wagstaff & Wheatcroft (2013) investigated verbal and nonverbal cues to deception utilizing thirty-two public appeals (sixteen deceptive and sixteen genuine).

The current research seeks to explore judgments of truthfulness made on full overall single accounts using real cases of false and genuine appeals for information, constituting ecological high-stake lying situations. That is, examine cases where a person makes a televised appeal for information regarding finding a relative that is missing, or for information that could lead to the detection of the person who murdered their relative, at a press conference attended by police officers and journalists or via a television or reporter interview about the disappearance/murder. By genuine appeal we mean that the appealer had been found not to be involved in the disappearance or murder of their family member and by false appeal that the appealer had been found to be the perpetrator of the murder of their family member. This
categorisation was based on the established ground truth as determined by a conviction in court.

In addition, most of the existing research has used college students or police officers. In the present experiment participants were academics from the disciplines of psychology, criminology and behavioural sciences. To the present researchers’ knowledge only one previous study (Ekman, et al., 1999) explored the lie detection performance of academic psychologists. But no study has explored this particular group in relation to judgments of truthfulness with the naturally occurring high stakes lies in false appeals. This is an interesting group, not only because participants would be familiar with lie detection research and have some understanding of the difficulties inherent within it but also because of the systematic way academics routinely assimilate information and strive to deconstruct initial assumptions of the veracity or deception of the appealer. Systematic processing would be more likely to occur from this participant group than from other groups such undergraduates or police officers the latter of which have been identified to hold a general communication suspicion (Levine & McCormack, 1991), which could increase their attribution of lying.

The Scale

Since so few studies have adopted a person perception approach to the study of judging truthfulness particular cues had to be identified from studies of deception as the indicators that people actually use to make veracity judgements.

The control approach or motivational impairment effect (DePaulo & Kirkendol, 1989) posits that liars are aware of the stereotypical markers of deception and make a concerted effort to control their behaviour for example maintaining eye contact and decreasing movement to
present a credible impression. However, this may be undertaken excessively, arguably because of a lack of awareness about the non-verbal behaviour actually displayed (Vrij, Edward, & Bull, 2001). Though, stereotypical gaze avoidance and excessive movements generally associated with lying behaviour are an unsafe basis to diagnose deceit. Alternatively, Ekman (1985/2009) views deception as an emotional event; the guilt of deceit, fear of detection or delight from duping the receiver produces emotional responses in the deceiver which manifest in their (non)verbal behaviour.

DePaulo et al.’s (2003) meta-analysis of studies focusing on emotional, verbal, facial and physiological cues to deception concluded that all these cues require further evidence for them to be regarded validly indicative of deceit. For DePaulo et al. discriminant validity is the fundamental weakness because cues regarded indicative of deception could similarly be indicative of the anxiety induced through fear of erroneously being accused of deceit. Further specific stereotypical cues of deception, such as gaze aversion and postural shifts, showed small overall effects, countering the cultural perception of being cues of deceit (DePaulo et al., 2003).

It appears that stereotypical, cultural markers that individuals regard indicative of deception, such as gaze aversion, are not reliable, availing the conclusion that there is little support for a universal behavioural signature of deception. Research has consistently identified that relying upon these behavioural indices to diagnose deceit provides an unsafe basis (Malone, DePaulo, Adams, & Cooper, 1997; Anderson, DePaulo, Ansfield, Tickle & Green, 1999). However, one area that lie detection research has recently begun to explore is implicit lie detection. That is, while observers may not be accurately able to diagnose deceit per se, verbalising thought processes pertaining to stimuli participants are viewing often elicits their
subtle scepticism when the message is deceitful rather than truthful (Hurd & Noller, 1988). DePaulo, et al’s (1997) meta-analysis lends support to implicit lie detection, identifying disparities in confidence ratings, with participants feeling more confident about their decisions with truthful than deceptive messages.

In their exploration of direct and indirect measures of deception Vrij, Edward & Bull (2001) found that police officers were not able to distinguish between truths and lies using a direct method (truth-lie), but only by using indirect methods of deceit detection. These results have been supported by Hurd & Noller (1988) and Anderson, DePaulo & Ansfield (1999) who all concur that the indirect measures of deception inspired greater accuracy in their participants than did the direct measures of deception.

Koper & Sahlman (1991) assessed the verbal and nonverbal correlates of naturally-occurring, high motivation deceptive communication. Categories of deceptive behaviour could be differentiated into strategic (i.e. vocal certainty, verbal directness, involvement, facial animation) and leakage cues (i.e. arousal, tension, facial pleasantness, verbal consistency, verbal plausibility).

In their research they found that perceived arousal was the single best predictor for distinguishing between liars and truthtellers. Liars were more aroused than truthtellers. Arousal, as a feature of deception is supported by other studies. Ekman and Friesen (1969) also refer to the expression of arousal indicators during deception as leakage. Ekman and Friesen’s (1969) leakage hypothesis is also supported by Ekman & Friesen (1974). Perceived tension was the second best predictor for distinguishing between liars and truthtellers. Liars were more tense than truthtellers. Perceived nervousness had the dominant impact/
strongest predictor on judgments of deception in Kraut & Poe's study (1980), was a strong predictor for Krauss et al (1976) and has been used by many cultures as part of a test for deception (Smith, 1967).

Verbal consistency and verbal plausibility may also be strong predictors on judgments of truthfulness. Advanced constructing deceptive accounts are more cognitively demanding than truth-telling (Walczyk et al, 2003; Vrij, 2008) therefore deceptive accounts are expected to demonstrate lower verbal consistency throughout the appeal, to the extent that what is said would be regarded less consistent. Further the perceived plausibility of the account, specifically in relation to the likelihood that the events could have transpired in the way in which the appealer presents it, would be hypothesised indices of truthfulness across appeals.

Koper & Sahlman (1991) found that the truth sounded more plausible than did deception, liars stories were perceived as less plausible and less consistent. Kraut (1978) examined what cues influence people's judgments about whether an actor was lying. Among other results they found that observers seemed to use the plausibility of an answer and well as internal consistency as cues to base their judgments on, a finding supported by other correlational and experimental studies (Apple, Streeter & Krauss, 1979; Kraut, 1978; Krauss, Geller & Olson, 1976), concluding that verbal consistency and verbal plausibility may be associated with either the truth of an account or observers' judgments of its truthfulness (Kraut, 1978).

Involvement, verbal directness and vocal certainty can be classified as strategic clues or leakage cues (Koper & Sahlman, 1991), to the extent to which they are intentionally displayed to create a vague account, which may prove functional for plausible deniability in the event of their deception being uncovered, i.e. by disputing the intention behind what was
said or the meaning behind it. Because deception is a cognitively complex task, indicators of this cognitive-load proliferate the liar’s behaviour including decreased movement, reduced attention to non-verbal behaviour (Mann, Vrij, & Bull, 2002) and increased reaction times to questions when the answer is deceptive (Walczyk et al., 2003). Additionally this may be communicated through increased speech disturbances and decreased speech rate (Vrij & Heaven, 1999) which observers may perceive as low vocal certainty and verbal indirectness. Although the caveat here is that intelligence, particularly verbal intelligence may negate the effects of the cognitive load on lying behaviour because of the ability to multitask and spontaneously fabricate plausibly (Kashy & DePaulo, 1996; Ekman & Frank, 1993).

Koper & Sahlman (1991) found that liars were less certain (vocal cues during the deceptive messages were rated as less certain, less direct and more involved in their presentation than truthtellers. Therefore truthful appealers would appear more verbally direct and more vocally certain since they are telling the truth about the transpired events.

As mentioned earlier facial appearance has a big impact on social perception and our judgements/impressions of others. Ekman and Friesen (1974) found that trained observers could make accurate judgements from observation of the face and Masip, Garrido & Herrero (2003) asked participants to assess the credibility of truthful and deceptive written messages based on facial appearance. The facial appearance of the purported senders of these written messages was conceptualised as a static cue that could yield judgments of truthfulness or deceptiveness. They found that when the statements were accompanied by babyfaced pictures, participants tended to judge them as truthful. Facial pleasantness and facial animation may therefore act as predictors for judgements of truthfulness. Participants’ perceptions of facial pleasantness may appear a misnomer in relation to appeals for
information, following Koper & Sahlman (1991) operationalisation as fewer smiles and so on. In the present study pleasantness is measured on the basis of how concordant facial expressions are with verbal and/or nonverbal behaviour of the appealer. Apparent discord would result in the appealer being perceived less facially pleasant and therefore less truthful. Although as, facial attractiveness does influence social judgements (Zebrowitz, 1997) there is the potential that appearance may influence creditability judgements (Masip, Garrido & Herrero, 2003) but in aesthetic terms, relating to demeanour-bias where particular faces are interpreted as particularly innocent or guilty looking (Bond & Robinson, 1988).

Taking into account the above research and the fact that the stimulus in this study closely matched that of Koper and Sahlam (1991), the following leakage and strategic cues were regarded appropriate to use in measuring participants’ perception of the appealers’ truthfulness: Facial pleasantness, facial animation, arousal, tension, involvement, verbal consistency, verbal plausibility, verbal directness and vocal certainty.

In addition to the strategic and leakage cues five variables associated with person perception of truthfulness were used in the current study: open, deceptive, genuine, trustworthy and emotional. By identifying qualities which are associated with perceptions of truthfulness; openness, non-deception, genuineness, trustworthiness and emotionality in addition to cues associated with deception the study intends to explore which cues and qualities are significantly associated with perceptions of truthfulness and predict the veracity judgments in the current study.
METHOD

Participants

Participants were thirty academics from a University in Northern England of whom 50% were male (N=15) and 50% female (N=15). The largest group were from the discipline of Psychology (70%), followed by Criminology (23.33%) and Behavioural Sciences (6.67%). The mean age was \(M=37.27\) years (SD=9.57; range=24-61 years). Twenty-eight participants (93.3%) were white and two participants (6.7%) Asian. The nationalities of the participants were: British (76.7%; N=13 males and 10 females), Irish (6.7%; N=1 male and 1 female), Swedish (3.3%; N= 1 female), Dutch (3.3%; N= 1 female), New Zealander (3.3%; N= 1 female), Greek (3.3%; N= 1 male) and American (3.3%; N= 1 female).

Materials

Participants saw twelve video clips of press conference appeals and news broadcast interviews. These were collected through internet broadcast sites, national and local news websites and the personal collection of Professor Canter. The collection of appeals were categorised as either genuine (where the appealer had been found not to be involved in the disappearance or murder of their family member) or false appeals (where the appealer had been found to be the perpetrator of the murder of their family member). This categorisation was based on the established ground truth as determined by a conviction in court. The cases in which no perpetrator had been identified, or where the investigation was ongoing were rejected from the final set.

Taking into account the inherent difficulties in precisely matching the genuine and false appeals the remaining false appeals were matched as closely as possible with the remaining genuine appeals in terms of the following criteria: a. The context of the appeal (press
conference, interview with news reporter in a studio or interview with a news reporter at the appealer's home), b. the sex of the appealer, c. the age of the appealer and d. whether other family members or police officers were present at the appeal. This was undertaken because of the potential impact of the context, sex, age and the presence and status of other people on the verbal and nonverbal behaviour of the appealer. In addition, as far as it was possible the appeals were additionally matched in terms of the relationship between the appealer and the missing/murdered family member, for example father and son or husband and wife.

From the cohort of matched genuine and false appeals only six pairs were retained for the purpose of experimental stimuli. In eliminating appeals consideration was given to the fact that some of the cases received considerable media attention and therefore it would be very likely that the participants would recognise the appealer or know the outcome of the investigation. The final stimulus consisted of twelve cases from England, USA, Australia, New Zealand and Ireland.

The content of the appeal consisted mainly of the appealer appealing directly to the ‘audience’ for information that would lead to the apprehension of the murderer (in cases where it was known that the person who had disappeared was actually murdered) or for information regarding the whereabouts of their family member (in cases where the investigation was still exploring the possibility of a disappearance). In some appeals there was an additional direct appeal to the disappeared person, specifically for them to contact someone and confirm their safety. In preparing the final format of the clips, during the editing process, questions from reporters that asked the appealer whether they had murdered their relative were removed because emotive words and questions such as ‘did you kill your wife?’ potentially elicit the activation of cognitions associated with deviancy/criminality.
(Neely, 1977). In addition, other questions contained in the appeals that were not directly relevant to the disappeared/murdered individual or the appealer were removed, however only if this could be done without impacting on the integrity of the clip, fundamentally retaining an unbiased shortened version of the appeal. The final appeals varied in length from 34 seconds to 185 seconds with a mean length of 96 seconds per appeal. The combined length of the 12 appeals was 18 minutes.

The questions that the participants were asked to answer after each appeal consisted of three sections; the first section explored person perception global attributes associated with judgments of truthfulness of the appealer in the form of a 5 point Likert-scale. The items were judgments of the appealer on the characteristics of being open, deceptive, genuine, trustworthy and emotional.

The next section explored person perception behavioural attributes associated with judgments of truthfulness of the appealer using a semantic differential, 5 point-Likert format. Taking into account previous research and existing literature (i.e. Kraut, 1978; Kraut & Poe, 1980; Koper & Sahlman, 1991, Ekman’s leakage clues) and the fact that the stimulus in this study closely matched that of Koper and Sahlam (1991), the following leakage and strategic cues were regarded appropriate to use in measuring participants’ perception of the appealer’s truthfulness: facial pleasantness (unpleasant/pleasant), facial animation (low/high), arousal (cool/bothered), tension (relaxed/tense), involvement (withdrawn/involved), verbal consistency (inconsistent/consistent), verbal plausibility (implausible/plausible), verbal directness (evasive/direct) and vocal certainty (uncertain/certain)
The final section explored participants’ perceptions of whether the appealer was telling the truth or not, again in the form of a 5-point Likert-scale with responses ranging from 1 (‘I am sure they were lying’) to 5 (‘I am sure they were telling the truth’) with 3 being unsure.

Procedure

The study took place at a University computer laboratory. Participants were asked to take part in an experiment concerning perceptions of public appeals for information regarding either the disappearance of a relative or for information to lead to the apprehension of the perpetrator who murdered the relative. Participants were tested individually although on a few occasions two participants took part in the experiment simultaneously, however were unable to see or hear which appeal the other person was watching.

In order to create a completely randomised and counterbalanced order of stimulus display 50 random sequences of 1-12 were created. Participants were then assigned a randomly sequenced slide show presentation using PowerPoint depicting the 12 appeals. They were seated in front of a computer, and further to standardised participant verbatim, which included practical direction in using the slide show format, were briefed as to the purpose of the study. The participants viewed the appeals on the computer screen wearing headphones. Participants began the first appeal by double clicking the mouse cursor on the starting frame, which was shown in a maximised full screen format. After watching each appeal the slide show stopped on the last frame of the appeal, at which point the participant then completed the questions pertaining to that appeal. When the participants were ready to watch the next appeal they clicked on the next starting frame and so on. The experiments took approximately an hour to complete.
RESULTS

To prepare data for analysis, cases, where participants had disclosed that they recognised the appealer or that they were familiar with the investigation to which an appeal was related to, were excluded. In all 11 such cases were omitted leaving 349 cases for further analysis.

Truth accuracy

Seventy-six out of a possible 176 true cases were correctly identified, providing a 43.18% accuracy total for detecting truthful appealers. However what has not been identified in previous literature, which this current study has explored, is that when people are not coerced into making a judgment of truthfulness, and can elect to state they are unsure as to the veracity of the appealer, a significant minority will indeed indicate they are unsure. When removing the unsure responses the overall truth accuracy rate was 55%.

Scale item correlations with the truth variable

Analysis was conducted to determine how judgments of truthfulness were made, across participants and across appeals. In doing this the data was explored to ensure that the scores were in the same direction, in that the higher the score the more likely it was to be associated with truthfulness. While this would appear to be a simple observation to make, as was the case for the deception scale item, 1 (not at all) – 5 (yes very), this was not necessarily the case for other semantic differential items, in particular arousal: (1=cool–5=bothered) and tension: (1=tense –5=relaxed). In Koper & Sahlman’s (1991) study liars were perceived significantly more tense and aroused than truth-tellers. However, because of the nature of the appeals for information in the current study it was less clear whether tension would be as linearly related to deception. Whether the observer would expect that the appealer would be in a very anxious, tense state following either the murder or disappearance of their family
member and thus judge this anxiousness and tenseness to be indicative of deception or truth, was an issue necessitating exploration.

In testing this notion the Likert-scale for the variable tension (1=tense, 2=quite tense, 3=unsure, 4=quite relaxed, 5=relaxed) was reversed, but retained for the anxious (1=cool, 2=quite cool, 3=unsure, 4=quite bothered, 5=bothered) in the content questionnaire. This permitted exploration of participant expectations of truthful appealers’ behaviour. The researchers hypothesised that perceived tenseness and anxiousness in an appeal for information would be regarded as adding veracity to the appealers account. Pearson’s correlations were run separately on the appeals in order to test which of the items required recoding. As was expected consistent negative correlations were found for the deceptive Likert-scale item across appeals and, additionally for the semantic differential tension. Accordingly these two variables only, were recoded, to reverse the scales. The correlations with the truth variable across aggregate appeals after recoding the variables deception and tension are shown below (Table 1).

| INSERT TABLE 1 HERE |

The variables with the highest correlation were, in descending order: genuine, trustworthy, verbal plausibility, open and deceptive, indicating that high scores on these variables were most positively correlated with high scores on the final veracity judgement. While all other variables (except tension) had a significant correlation ($p<.01$) the correlations were not above .816 indicating more of a moderate relationship between these variables and the truth judgement. Notably the variable tension had a very low correlation at $r =105$, $p>.05$ indicatively of a minimal association with the overall truth judgement. Nevertheless the table
supports the hypothesis that there is an underlying significant positive association with perceptions of truthfulness apparent in the items of the scale, with the exception of the item tension.

**Judgments of truth across all appeals**

Further analyses were undertaken to identify those items most strongly predictive of a truth judgment. Multiple regression analyses were conducted with the truth judgment serving as the criterion variable and fourteen variables as predictors. The multiple regressions across appeals are presented below (Table 2).

The analysis yielded a significant multiple correlation ($R^2=.763$, $F(16, 332)=66.67$, $p<.001$). The variables of openness, (non)deceptiveness, genuineness, trustworthiness and verbal plausibility were significant predictors of the criterion truth judgment across participants and all 12 appeals, indicating how the general truth judgments are made across the sample. Consequently perceptions of these five variables give rise to the best prediction of the criterion variable, while the remaining variables (emotional, facial pleasantness, facial animation, arousal, tension, involvement, verbal consistency, verbal directness and vocal certainty) do not add any significant predictive power to the model.

The $R^2$ statistic shows that within the sample 76.3% of the variance is accounted for by this model. The adjusted $R^2$ statistics, while being a more conservative estimate of the amount of variance accounted for in the model, still indicates that 75.1% of the variance is accounted for and that the model has strong predictive power, in relation to the general population. From
the above table it can be seen that while all the variables have a significant impact on the truth judgment the predictor variable *genuine* has the greatest impact at .300, indicating the estimated standard deviations rate of change upon the criterion variable will be .300 by a change of one standard deviation of this predictor variable. The next variable with the greatest estimated effect on the criterion truth variable is the predictor *verbal plausibility* with .205 effect for every one standard deviation change upon the predictor. The variable with the least significant effect on the criterion is *open* with a .112 standard deviation effect on the criterion.

As Pearson’s correlation showed the variables *genuine* and *trustworthy* had correlation coefficients of (.816 and .784) with the judgment of truth, the possibility of the collinearity of these two variables was taken into account. However this possibility was eliminated by examining the Variance Inflation Factor (VIF) which, for each variable, was between 1.14-4.82. For genuine and trustworthy it was 4.82 and 4.06 respectively. It has been noted by some authors that VIP>10 shows strong evidence of collinearity which would then effect the regression efficient indicating poor estimation of the criterion (Kinnear & Gray, 2010). Further the Durbin-Watson statistic=2.03 indicates a weak negative correlation eliminating the possibility of collinearity. The Durbin-Watson statistic tests for the correlations between errors, whether adjacent residuals are correlated; one assumption in the regression analysis is that the residuals are independent (Field, 2009).

The regression casewise diagnostics showed that one case in the data set had a large standard residual values (-3.03) indicating the possible infringement of the assumptions of the analysis potentially indicating the validity of the results may be in doubt (Field, 2009). However the histogram of the residuals showed relatively normal distribution mean=2.55 (SD=.98)
additionally the mean standardised residual was .000 and Cook’s Distance was .003 indicating the assumptions were indeed met and the results drawn can be considered valid.

This regression analysis shows, across both the deceptive and genuine appeals, that participants have assigned their diagnosis of the appealer’s truthfulness on the basis of how open, (non)deceptive, genuine, trustworthy and verbally plausible they perceive the appealers to be.

**Judgments of truth across false appeals**

Further analysis was conducted on the false appeals only (Table 3).

<table>
<thead>
<tr>
<th>INSERT TABLE 3 HERE</th>
</tr>
</thead>
</table>

The analysis yielded a significant multiple correlation ($R^2=.771$, $F (16, 156)=32.811$, $p<.001$). The variables of openness, genuineness, trustworthiness and verbal plausibility were significant predictors of the criterion truth judgment across participants and the six false appeals. Emotional, deceptive, facial pleasantness, facial animation, arousal, tension, involvement, verbal consistency, verbal directness, vocal certainty were not significant predictors in this model.

Therefore deceptive appealers were regarded truthful, if participants perceived them as significantly more open, genuine, trustworthy and verbally plausible. The variable with the greatest impact upon this model is the perception of trustworthiness with a standardised beta coefficient of .254, indicating a predicted .254 change in the standard deviation for every one standard deviation positive increment in the criterion variable. The variable with the lowest significant effect is openness with .162 standard deviation increase with every one for the
criterion. In this model the $R^2$ statistic shows that 77.1% of the variance is accounted for within this sample. The adjusted $R^2$ statistics provides shrinkage of 2.4% indicating that the amount of variance accounted for within the general population would be 74.7% showing that this model can be regarded as having strong predictive power.

**Judgments of truth across genuine appeals**

Further analysis was conducted on the genuine appeals only (Table 4).

| INSERT TABLE 4 HERE |

The analysis yielded a significant multiple correlation ($R^2=.790$, $F$ (16, 159)=37.420, $p<.001$). From this third model the predictor variables deceptive, genuine, tension and verbal plausibility were significant predictors of the criterion variable truth, while the remainder variables (open, trustworthy, emotional, facial pleasantness, facial animation, arousal, involvement, verbal consistency, verbal directness and vocal certainty) do not have any significant predictive power. As the two variables deceptive and tension have been recoded, higher scores identify the appealer as perceived (non)deceptive and tenser. Consequently appealers perceived less deceptive and more genuine, tense and verbally plausible were significantly more likely to be judged truthful within this sample. The predictor variable with the greatest impact upon this model was genuineness (standardised Beta coefficient .363) and the lowest was tension with .114 indicating for every single positive standard deviation increment in these predictors would cause .363 and .114 respective increase in the criterion truth. The $R^2$ statistic indicates that 79% of the variance on the criterion is accounted for by the predictors within this sample. The adjusted $R^2$ statistics provides shrinkage of 2.1%
indicating that the amount of variance accounted for within the general population would be 76.9%, again indicating this model has strong predictive power.

Interestingly the items tension and deception in genuine appealers were found to be predictive of the criterion variable truthful. While for the false/deceitful appealers the predictor variables tension and deception were not predictive. However for the false/deceptive appealers predictor variables trustworthy and open were found predictive of the truth judgment.

DISCUSSION

The current study explored how judgments of truthfulness are made. Unlike past work, this study explored the person perception correlates of truthfulness as it is perceived by the observer using thirty academics who were asked to judge the truthfulness of appealers during an appeal. Perceptions of global person characteristics, notably openness, non-deceptiveness, genuineness, trustworthiness were related to judgements of truthfulness. The indications are that person perception is indeed relevant to truthfulness judgment. The identification of some of the specific attributes that are pertinent to these judgements now opens the way for exploration of which of these person perception qualities observers rely upon to make their judgements are valid indicators of accurate truth judgements.

Although these are not unexpected, identifying the particular perceived attributes of the person that were the basis for judgements opens up potential research directions in truthfulness detection. The strongest predictors of the truth judgement were the global person perception indices of openness, (non)deceptiveness, genuineness, trustworthiness and the behavioural measure of verbal plausibility. This suggests that judgements are based upon a
complex interplay of personality assessments, such as general openness and trustworthiness, and the plausible presentation of the account. It seems likely that a plausible presentation of the account will relate to its content; an issue that requires further study.

There were only small variations in the predictive person perception variables across the genuine and false appeals. Interestingly, for deceptive appeals perceived truthful over-attribution was present. Further regression analysis was performed on the deceptive appeals alone and identified similar significant predictors; open, genuine, trustworthy and verbal plausibility. The multiple regression conducted upon the genuine appeals alone yielded the predictor variables, (non)deceptive, genuine, tension and verbal plausibility.

It is interesting that perceptions of non-deceptiveness were significant predictors in the genuine appeals and not for the false appeals. This was regardless of the accuracy of judgments especially in relation to the fact that perceptions of genuineness were significant positive predictors in both the false and truthful appeals. It could be inferred that the variable (non)deceptiveness is connected to more intuitive perceptions of veracity which were significantly impactful for the truthful appeals, but not for the false appeals. Participants appear more reluctant to regard the false appealers as non-deceptive than genuine appealers but were as equally willing to infer genuineness in both false and genuine appeals. Therefore, while genuineness was a significant predictor in both models it may have been an easier, less intuitive perception to make. Indeed perceiving someone as non-deceptive may be inferred as a leap of faith, as trust, to a certain extent, is invested in that particular perception to a greater extent than in perceptions of genuineness. Additionally, as the truth-bias is often accompanied by over-attribution it could be inferred that although a truth-bias was not found in the present study for deceptive appeals remnants of over-attribution were.
The significant predictor tension in the genuine appeals is also interesting. Lie detection literature has identified that deceptive individuals often are perceived as tenser than truth-tellers. Indeed Ekman (2009) advances that lying is an emotional event where guilt about lying or fear of being caught often manifests in tense behaviour. However the present study found that tension is correlated with perceptions of truthfulness. Indicatively it is expected that bereft relatives, or those appealing for information regarding the whereabouts of a disappeared relative, would appear tense, but also because they are in an unfamiliar, intimidating environment where they required to talk publicly while being filmed.

Furthermore because tension was a significant predictor of truthfulness for genuine and not false appealers this is again arguably an intuitive perception. Deceptive appealers would also be likely to appear tense in this setting as the stakes are so high. However the control approach (DePaulo & Kirkendol, 1989) advances deceptive individuals are aware of these stereotypical markers of deception and make a concerted effort to subvert this affect to present a credible impression, potentially accounting for the absence of this item in the deceptive appeals regression. Verbal plausibility in addition to genuine was a significant predictor in all three regression analyses. The significance of verbal plausibility is commonsensical as the participants were exposed to both the visual and verbal modalities of the appeal therefore perceptions of verbal plausibility is bound to factor high in the perceptions of truthfulness (Burgoon et al., 2008). Furthermore regardless of the ultimate veracity of the appealers their versions of events had previously been retold to the police, thereby practiced, and to a great extent, accepted by the investigating officers as plausible, indicatively observers perceptions of verbal plausibility is a logical finding.
The present study provides this exploration of perceptions of truthfulness in naturally occurring high-stake situations implying strong ecological validity. Criticisms outlined by Koper & Sahlman (1991) regarding typical lie-detection studies pertaining the laboratory’s non-conducive facilitation of high-stake deceptive behaviour, the induction of what is crudely participant’s deceptive performance and the fundamental omission of the crux of deception, self-serving personal agency are all factors constituting the strengths within this study.

One potential limitation in the present study is that the length of the clips for each appealer was between 34-185 seconds. It could be argued that this range gave an unfair advantage to detection of truthfulness in appeals that were longer, although the accuracy results do not necessarily support this proposition. Further as Feeley & deTurk, (1995) identify veracity judgments can be made quickly in interactions and with little behavioural information therefore this was not viewed as problematic. Moreover as Vrij & Mann (2001a) utilised relatively brief appeals mean=42seconds (range=22-62seconds) the duration of present appeals was regarded acceptable.

One particular issue that emerged pertained to the item facial pleasantness. It was found that some participants interpreted this item as exploring attractiveness, while others, how concordant appealer’s facial expressions were with the conveyed message. While the present researchers intended the item to pertain to the latter interpretation, to a certain extent, this is not as problematic as it appears. Perceptions of attractiveness/pleasantness are in wholly subjective and arguably instinctive. Therefore participants who thought their assumptions were made on more aesthetic, attractiveness grounds may still have been affected subconsciously by their implicit veracity assumptions.
In the present study features of the appealers’ openness, non-deceptiveness, trustworthiness and genuineness may have encapsulated both verbal and bodily behaviour of the appealer in observer’s veridical judgment. However, unlike Feeley & deTurk (1995), the current study did not utilise an open question from which responses were coded into respective categories, therefore findings do not allow exploration of this possibility. Future research may seek to explore the psychometric properties of, for example, perceived verbal and nonverbal openness, deception, trustworthiness and genuineness and use these as distinct categories to find which remain significantly associated with the veracity judgment. Indeed existing literature (c.f Burgoon et al. 2008; Vrij & Mann 2001b) omitted exploration of the psychometric properties of person or behavioural perception as an exploratory issue in their studies. However in pursuit of holistically understanding the perceptual correlates of truthfulness in high stake deceptions future research needs to attend to this issue.

In expounding this area of research it is apparent that the present study only explores participant’s perceptions of the appealer’s truthfulness. Consequently more questions arise in relation to how the appealer’s particular communicative or narrative style elicits observer’s perceptions of truthfulness, requiring exploration. This would be a particular insightful investigation as it would explore whether particular narrative styles drawn on by the appealers are more effective at eliciting erroneous or correct inference of truthfulness.

Most lies are relatively innocuous, with little more than sanctions of social disapproval if discovered. A minority of deceits are more insidious, such as lying during the course of police investigation. Here offenders may lie to evade further scrutiny or arrest; witnesses may withhold critical evidence, embellish or fabricate their versions of events (King & Dunn,
During an investigation and indeed upon its inception when the crime is reported, it would be invaluable to discern whether the reporter or witness is truthful.

This exploratory study examined ecologically valid cases of televised appeals for information regarding missing relatives, or for information that could lead to the detection of the person who murdered the appealers’ relative thus taking the first steps in understanding the correlates of deception and truthfulness detection. Consideration and development of these current findings in future research may ultimately provide investigating officers, faced with a murder or disappearance investigation, where the truthfulness of the witness or public appealer is in doubt, with a repertoire of research conclusions, which may have utility in the identification of truthful and deceptive appeals for information.
REFERENCES


Table 1. Scale item Pearson’s correlations as a function of the truth judgement across participants and appeals/the 349 cases.

<table>
<thead>
<tr>
<th>Scale Item</th>
<th>Truth</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>.717**</td>
<td>3.27 (1.21)</td>
</tr>
<tr>
<td>Deceptive</td>
<td>.700**</td>
<td>3.24 (1.19)</td>
</tr>
<tr>
<td>Genuine</td>
<td>.816**</td>
<td>3.21 (1.19)</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>.784**</td>
<td>3.06 (1.10)</td>
</tr>
<tr>
<td>Emotional</td>
<td>.427**</td>
<td>3.30 (1.43)</td>
</tr>
<tr>
<td>Facial Pleasantness</td>
<td>.421**</td>
<td>2.91 (.97)</td>
</tr>
<tr>
<td>Facial Animation</td>
<td>.313**</td>
<td>2.87 (1.26)</td>
</tr>
<tr>
<td>Arousal</td>
<td>.362**</td>
<td>3.22 (1.35)</td>
</tr>
<tr>
<td>Tension</td>
<td>.105</td>
<td>3.30 (1.27)</td>
</tr>
<tr>
<td>Involvement</td>
<td>.451**</td>
<td>3.36 (1.18)</td>
</tr>
<tr>
<td>Verbal Consistency</td>
<td>.519**</td>
<td>3.53 (1.14)</td>
</tr>
<tr>
<td>Verbal Plausibility</td>
<td>.715**</td>
<td>3.34 (1.18)</td>
</tr>
<tr>
<td>Verbal Directness</td>
<td>.571**</td>
<td>3.47 (1.17)</td>
</tr>
<tr>
<td>Vocal Certainty</td>
<td>.597**</td>
<td>3.02 (1.29)</td>
</tr>
</tbody>
</table>

Significance: * p<.05 ** p<.01 *** p<.001
Table 2. Significant variables across the total sample \((N = 349)\) over twelve appeals

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>b</th>
<th>SE b</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.976</td>
<td>180</td>
<td>.112*</td>
</tr>
<tr>
<td>Openness</td>
<td>.123</td>
<td>.049</td>
<td>.152***</td>
</tr>
<tr>
<td>Deceptive</td>
<td>.171</td>
<td>.045</td>
<td>.300***</td>
</tr>
<tr>
<td>Genuine</td>
<td>.338</td>
<td>.066</td>
<td>.151**</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>.186</td>
<td>.066</td>
<td>.205***</td>
</tr>
<tr>
<td>Verbal Plausibility</td>
<td>.233</td>
<td>.053</td>
<td>.151**</td>
</tr>
</tbody>
</table>

\(R^2 = .763 \) Adjusted \(R^2 = .751; F_{16,332} = 66.67***\)

Significance: * \(p<.05\) ** \(p<.01\) *** \(p<.001\)
Table 3. Significant variables across the total sample (N=173) over the six false appeals

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>b</th>
<th>SE b</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.074</td>
<td>.315</td>
<td>.162*</td>
</tr>
<tr>
<td>Openness</td>
<td>.181</td>
<td>.081</td>
<td>.221*</td>
</tr>
<tr>
<td>Genuine</td>
<td>.251</td>
<td>.101</td>
<td>.254**</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>.311</td>
<td>.102</td>
<td>.221***</td>
</tr>
<tr>
<td>Verbal Plausibility</td>
<td>.251</td>
<td>.076</td>
<td>.221***</td>
</tr>
</tbody>
</table>

$R^2 = .771$ Adjusted $R^2 = .747$; $F_{16,156} = 32.811^{***}$

Significance: * p<.05 ** p<.01 *** p<.001
Table 4. Significant variables across the total sample ($N=176$) over the six genuine appeals

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>b</th>
<th>SE b</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.936</td>
<td>.267</td>
<td>.222***</td>
</tr>
<tr>
<td>Deceptive</td>
<td>.239</td>
<td>.061</td>
<td>.363***</td>
</tr>
<tr>
<td>Genuine</td>
<td>.402</td>
<td>.087</td>
<td>.114**</td>
</tr>
<tr>
<td>Tension</td>
<td>.127</td>
<td>.046</td>
<td>.200**</td>
</tr>
<tr>
<td>Verbal Plausibility</td>
<td>.217</td>
<td>.080</td>
<td></td>
</tr>
</tbody>
</table>

$R^2 = .790$ Adjusted $R^2 = .769$; $F_{16,159} = 37.420^{***}$.

Significance: * $p<.05$ ** $p<.01$ *** $p<.001$