Improving professional observers' veracity judgments by tactical interviewing

Sandham, A, Dando, C.J., Bull, R. and Ormerod, T.

Improving professional observers' veracity judgments by tactical interviewing

Alexandra L. Sandham [1], Coral J. Dando [2], Ray Bull [3], Thomas C. Ormerod [4]

[1] University of Gloucestershire, UK

[2] University of Westminster, UK

[3] University of Derby, UK

[4] University of Sussex, UK

Abstract

Understanding whether a person of interest is being truthful during an investigative interview is a constant challenge and is of concern to numerous criminal justice professionals, most of whom are not involved in conducting the interview itself. Here we investigated police observers' veracity detection performance having viewed interviews with truthtellers and deceivers using either the Tactical Use of Evidence (TUE), Strategic Use of Evidence (SUE) or a Control technique. Thirty serving police officers participated as post interview observers and each viewed 12 interviews in a counterbalanced order. After each interview the officer made a veracity judgment. Overall, untrained police observers were significantly more accurate (68%) when making veracity judgments post TUE interviews, whereas for both SUE and Control performance was around chance (51% and 48%, respectively). Veracity performance for liars and truthtellers revealed a similar pattern of results (67% liars; 70% truthtellers) in the TUE condition. These results lend further support to the psychological literature highlighting the importance of how and when to reveal evidence or any other relevant event information during an investigative interview for 'outing' deceivers as well as allowing truthtellers early opportunities to demonstrate their innocence.

Keywords: Police Observers; Tactical Interviewing; Detecting Deception

1. Introduction

Understanding whether a person of interest is being truthful during an investigative interview is a constant challenge. Whilst the psychological literature highlights differences in the behaviour of deceivers and truthtellers, making judgments about whether a person is dissembling or feigning innocence in everyday legal settings is difficult. Reliable cues are not readily discernible and individual differences mean that cues presented are often inconsistent and misunderstood and are often not even important. Further, where legal requirements dictate the disclosure of evidence/information prior to interview (as is the case in the UK, for example), guilty persons of interest have ample opportunity to devise and practice dissembling scripts and behaviours prior to any formal interview to appear convincing, which introduces additional challenges for those making the "truth" or "lie" judgement (Anolli & Ciceri, 1997; Hartwig et al., 2007; Leins et al., 2013; Sporer, 2016; Verigin et al., 2019).

There is a body of academic literature concerned with improving police interviewers' ability to detect when an interviewee is being veridical (Dando & Bull, 2011; Dando et al., 2015; Dando & Ormerod, 2019; DePaulo et al., 2003; Ormerod & Dando, 2014; Vrij, 2019). However, whether a person of interest is dissembling is of concern to numerous other criminal justice professionals, most of whom are not involved in conducting the interview itself. Rather, they are professional observers tasked with considering the information emanating from the interview or the conduct of the interviewee during the interview. In the UK, for example, decisions concerning the likelihood of a successful prosecution are made post-interview by professional observers such as solicitors and barristers employed by the Crown Prosecution Service. Equally, in the case of serious crime (e.g., murder; terrorism) persons of interest are typically interviewed on numerous occasions. Here, professional police interview advisors are tasked with supporting interviewing officers with developing interview tactics and in doing so they often observe ongoing interviews or previously

conducted interviews. Therefore, investigating how to support professional observers in making judgments about dissembling and understanding how they make judgments is important because this can have significant ramifications for the processes of criminal justice.

Various psychologically guided methods for amplifying indicators of deceit during investigative interviews have been developed and empirically evaluated. One technique is the Strategic Use of Evidence (SUE). SUE is based upon two premises, the first being that there are differences in the cognition of deceivers and truthtellers, and second that both deceivers and truthtellers try to control their self-presentation in order to convince an interviewer that they are being truthful (Hartwig et al., 2010). The SUE technique requires interviewers to withhold evidence or event information from the interviewee to assess interviewee strategies when answering specific questions (Granhag & Hartwig, 2008). In brief, SUE interviews begin with a free account where interviewees are asked to provide details of their involvement in an event uninterrupted by the interviewer. The interviewer then explores the freely provided account and asks specific questions related to evidence or event information they are aware of without revealing what that evidence or information is. The aim is to identify whether the interviewee is forthcoming and truthful in their answer (i.e., is willing and able to provide more details) or employs an avoidance and/or denial strategy (i.e., offers scant details, skirting around the question or simply denying knowledge).

A consistent body of empirical findings reveals the efficacy of the SUE technique for improving the accuracy of veracity judgments (Hartwig et al., 2006; Luke et al., 2013; Luke, Hartwig, Joseph, et al., 2016; Luke, Hartwig, Shamash, et al., 2016). Police interviewers trained in the SUE interview technique were found to be more accurate when judging the veracity of the interviewee than officers who used their usual interview technique (65% vs. 43% accurate) (Luke, Hartwig, Joseph, et al., 2016). Typically, police interviewers disclose the evidence or information known to them and why they believe it to be incriminating at the

start of an interview (Bull, 2014). We refer to this as an Early interview, that is where the information is disclosed at the start, and in bulk. Similar results have been reported for untrained student lay observers who were more accurate in their veracity judgments when observing SUE interviews than Early interviews (61.7% v 42.9%) (Hartwig et al., 2005). However, as far as we are aware, no research has investigated whether the SUE technique improves the performance of professional observers such as police or other professionals in the criminal justice system who have not been trained in the SUE technique.

A second technique in the psychological literature is the Tactical Use of Evidence (TUE) technique (Dando & Bull, 2011; Dando et al., 2015; Ormerod & Dando, 2014; Parkhouse & Ormerod, 2018). TUE uses an incremental approach to information revelation whereby interviewees are asked to account for/explain each piece of information/evidence piece by piece. Where appropriate, interviewees are immediately challenged if their responses to questions are inconsistent or at odds with known information (be this incriminating or not). Otherwise, responses are immediately acknowledged and accepted as correct and consistent as the interview progresses. TUE was designed to limit opportunities for verbal maneuvering because interviewees have to account for each piece of information, individually and prior to being alerted to its nature. Deceivers become encircled in such a manner that weaknesses/discrepancies in their verbal armature are highlighted from the outset. Conversely, by its very nature TUE offers truthtellers the opportunity to immediately self-correct and/or explain responses to each question as necessary thereby alleviating any concerns about the veracity of their accounts in a timely manner. As with SUE, no research has investigated whether the TUE technique improves the performance of professional observers who have not been trained in the technique.

An empirical test of TUE and SUE techniques compared with the Early disclosure of evidence with police investigators trained in each technique, found a significant advantage

for detecting both deceivers and truthtellers using TUE (67% and 74% accuracy, respectively) versus SUE (54% and 42%) and early disclosure (53% and 47%) interviews (Dando & Bull, 2011). Officers were significantly more confident in their judgments following TUE interviews than both SUE and Early. Furthermore, when untrained laypersons viewed the three types of recorded interviews (Dando et al., 2015) there was a significant advantage for the TUE technique, with deceiver and truthteller accuracy of 66% for deceivers and 76% for truthtellers versus SUE (deceivers 54% vs truthtellers 44%) and early disclosure (deceivers 50% and truthtellers 48%). Lay observer's deception detection performance improved by 16% compared with the traditional Early interviews, whereas performance gains were modest (4%) in the SUE condition. Lay observer's veracity judgments were far stronger following a TUE interview than in both the SUE and Early, and deceptive interviewees reported finding TUE interviews far more demanding than both SUE and Early interviews (Dando et al., 2015).

We believe this pattern of results emerged because the cognitive load faced by interviewers using SUE was raised; as they had to maintain a complete account of the interviewee's story and the evidence to be revealed throughout the interview. SUE resulted in increased false positive judgments against truthteller interviewees as where evidence was revealed at the end of the interview, truthtellers were unable to respond to a barrage of evidence presented after they had provided what they deemed to be a convincing account. In contrast, TUE allows the interviewer to present evidence at appropriate moments incrementally throughout the interview, which both lowers the interviewer's cognitive load and also allows the truthful interviewee to respond to new evidence in a managed way. Thus, like SUE, TUE interviewing seeks to disrupt a liar's scripts and exploit gaps in their account (Levine, 2014). However, TUE confers advantages because it provides innocent interviewees early opportunities to convey their honesty and supports interviewer cognition by

highlighting the investigative value of the available information on a piece-by-piece basis, more quickly than the SUE technique (Dando & Bull, 2011; Dando et al., 2015). Since police interviewers (and observers) typically have to manage large amounts of event information, managing interviewer and observer cognition in the absence of immediate electronic assistance is an important consideration.

1.1 The current study

The research reported here explores professional police observers' veracity detection performance when viewing interviews with truthtellers and deceivers. We hypothesised that the timing of evidence release would assist police observers to make more accurate varacity judgements. SUE and the Early (control) release of evidence both involve the bulk release of evidence but at different stages of an interview. Release of evidence en masse, as occurs in Early (at the beginning) and SUE (towards the end) gives truthtellers little chance to recover (in the case of SUE) and alerts liars to what is known (Early) and may exacerbate indicators of deception among both liars and truthtellers. The incremental release of evidence in TUE interviews should allow clearer separation of truthtellers and liars by observers because it allows truthtellers to prove their innocence.

In order to test this hypothesis, professional observers were presented with video recordings of interviewees being interviewed after taking part in game-based activity, the 'Dodgy Builders' task (Sandham et al., 2015), in which the players were assigned truthteller and liar roles. The truthtellers were tasked with being builders of an Olympic venue and were required to undertake a race to complete a simulated construction task first. The liars were tasked to pretend to be undertaking the same construction task, whilst undertaking a different task to undertake a simulated destruction of the same Olympic venue without being discovered. Each player developed their own game plan, thereby allowing liars to construct their own deceptions, an important enhancement over standard deception detection paradigms

where liars are instructed on what deceptions to use. After completion of the game, each player was interviewed by a professional interviewer (Dando & Bull, 2011; Dando et al., 2015; Sandham et al., 2011; Sandham et al., 2015), who was provided with five pieces of evidence taken from game play that could be used to test the accounts of each player. In the current study, observers viewed videos of these interviews, and were tasked with making veracity decisions about each player's interview performance. Specifically, whether the player (mock person of interest) was a truthteller or a liar, the scale of truth telling/lying and the level of confidence they have in their decision.

2.   Method

2.1 Participants

One hundred and fifty-one participants took part as mock persons of interest. There were 69 men and 82 women with a mean age of 21.3 years (SD = 4.56), ranging from 18 to 54 years. Thirty serving police officers participated as post interview observers. There were 23 men and 7 women with a mean age of 34.37 years (SD = 6.55) and a mean length of service of 11.70 years (SD = 6.23) all employed by four British police forces. All observers were blind to the aims of the research, research questions and interview conditions. All observers were PEACE trained interviewers, and all had completed the Professionalising Investigation Programme (PIP) level 1 and so were trained to interview victims, witnesses and suspects in relation to priority and volume crime investigations, only. PEACE is an acronym for the stages of an investigative interview (Planning and preparation; Engage and explain; Account; Clarification and challenge; Evaluate). PEACE is prevailing interview framework underpining the UK College of Policing professional practice (College of Policing).

Interviews were conducted by one interviewer (to limit interviewer variability) with

over 10 years' specialist police investigative interviewing experience. The interviewer completed additional training that explained in detail the three interview procedures (early, late, and gradual), how each procedure should be applied (using interview protocols and written examples), and the rules of the game. The game rules explained only the truthtellers' task, thereby increasing the realism of the investigative role. Information regarding the deceiver's task (those playing the role of terrorists) was withheld from the interviewer. The researchers held two interactive sessions with the interviewer during which the game and interview procedures were verbally explained. The interviewer undertook several practice interviews, and received feedback.

2.2 Procedure and Materials

The study comprised four distinct phases: i) game play, ii) interview, iii) post-interview questionnaire and iv) police observer judgment, as described below:

2.2.1 Game play. Participants played the Dodgy Builder game in groups of four. Upon arrival participants were individually greeted by the researcher, and randomly allocated to either the role of a builder (truth teller) or a terrorist (deceiver). They were each shown role specific training videos (viewed individually on separate laptops with headphones), which explained the game rules, their specific role and how to use of the software (Dando et al., 2015; Sandham et al., 2011; Sandham et al., 2015). Participants played the game on individual laptops but all participants were in the same room. The software generated random dice throws and the participant traversed the board visiting virtual shops in order to buy building materials for their tasks and delivering these materials to a virtual building site.

The builder's (truthteller) task was to build part of a virtual Olympic stadium whereas the terrorist's (deceiver) task was to blow up the stadium. The list of required items to complete each task differed slightly as a function of the two roles. Terrorists were given details of the builder's task to enable them to consider how to mask their true identity both

when purchasing the required items during the game, and to construct a deceptive account later during the post game interview.

The game was divided into phases. At the end of each phase a (covert) investigator was able to observe a number of items related to the game play in that phase, i) all items bought from the shops (but not which participants had bought what), and ii) the places visited by each participant (but not the route taken). This investigator was able to choose two of the four participants to conduct a virtual 'stop and search', consisting of weighing 1 virtual van and asking to view 2 items in another virtual van.

All participants were paid £23 to take part, an additional £18 following completion of the game (the latter amount being an incentive to complete the game), and a further £5 following completion of the interview (an incentive to carry out the researcher instructions to convince the interviewer that they were truthtellers). Regardless of interview outcome all participants received the final £5 payment.

2.2.2 Interview. All mock persons of interest were provided with the following pre-interview instructions: *'Your task is to convince the interviewer that you are a builder'*. No guidance was given as to how this was to be/could be done. To limit interviewer variability, all participants were interviewed by the same interviewer with over 10 years specialist police investigative interviewing experience.

Following each game, the interviewer was given 4 case files (one for each participant taking part in that particular game). Each case file comprised information about that participants game play: i) locations visited in each game phase, and ii) items purchased from the shops for each game phase. In addition, if that participant's van had been stopped and searched – the weight of the van and the items shown by the participants was also included. The mean number of information items presented to the interviewer in each case file was 11.7 ($SD = 1.09$), ranging from 8.81 to 13.09.

Control - Interviews commenced with an *introduction* and *explain phase* followed by the interviewer disclosing all pieces of information selected from the case file. The interview continued with a *free recall phase* and then a *questioning phase* where the pieces of information were revisited. The responses to the *questioning* phase were challenged appropriately until each piece of information was accounted for. The interview concluded with a *closure phase*. More detailed information for the interview conditions and results can be found in (Dando et al., 2015).

Strategic – Interviews in this condition commenced with an *introduction* and *explain phase* as above but these were immediately followed by a *free recall phase* followed by the *questioning phase* where questions concerning each of the pieces of information were asked (without revealing the information held or the source of information itself). The interview continued with the interviewer finally revealing all of the pieces of information. If necessary, the interviewer revisited and challenged explanations given by the participants until they were satisfied with the responses received. The interview concluded with the *closure phase*.

Tactical – The *introduction*, *explain* and *free recall phases* were the same as described for the late disclosure condition. Then the interviewer commenced the *questioning phase* by revealing the first piece of information from the case file, asking the participant to provide an explanation/account for this information. The explanation was challenged or accepted as appropriate until the interviewer was satisfied and then the next piece of information from the case file was addressed in the same manner until all of the information from the case file had been addressed to the interviewer's satisfaction. The interview concluded with a *closure phase*.

2.2.3 Post interview questionnaire. Each participant (mock person of interest) individually completed a short post interview questionnaire comprising just three questions, each collecting quantitative data using a Likert scale format for i) motivation (1 = *not*

motivated at all to 7 = *very motivated*), ii) deception (1 = *I was completely truthful* to 7 = everything I said as a lie), and iii) cognitive difficulty experienced during the interview (e.g. 1 = *not at all difficult* to 7 = *very difficult*).

2.2.4 Police observer questionnaire. Thirty serving police officers from 5 UK police forces individually viewed 12 video recordings of interviews (2 truthtellers and 2 liars from each interview condition) that were for each interviewer randomly selected but were matched for duration. Officers viewed the interviews at times and dates to suit them by accessing the videos via a one-time access link. Officers were sent links to the videos, one-by-one, to prevent observers making relative judgments, which may have occurred if they had watched the interviews in bulk and compared across interviews.

Immediately post viewing each interview observers completed a questionnaire comprising eight questions: i) four demographic questions (age, length of service, police force, gender), ii) three questions related to the observers' task – lie decision, lie scale and confidence scale. The former was a dichotomous forced choice (yes; No). The latter two used a Likert style format (1 = *definitely telling the truth/completely confident* to 7 = *definitely lying/ not at all confident*). The final question was an open-ended invitation to explain how they made their decision.

3. Results

3.1 Player Post Interview Perceptions

A series of two-way analysis of variance (ANOVA) were used to investigate the impact of our experimental manipulations (independent variables) on the dependent variables. ANOVA is an inferential statistical method for examining the main effects of each of the independent variables separately (participant veracity and interview condition) and the interactions between the two.

Bonferroni corrections were applied regarding the player's self-report motivation, deceptiveness and difficulty as a function of Veracity Group (builder; terrorist) and Condition (Control; Strategic; Tactical). Motivation to comply with the experimenter's pre-interview instructions was high (M $_{motivation}$ = 5.70, SD = .87) and no main effects nor the interaction were significant (see Table I for means and SDs), all Fs< 3.621 all ps > .029.

There was a significant main effect of player Veracity Group (builder; terrorist) for deceptiveness, $F(1, 145) = 243.116$, $p < .001$, η_p^2 .63. As expected,terrorist players reported being more deceptive than builder players. The main effect of Interview type was non-significant, $F = 1.445$, $p = .239$. There was a significant Veracity Group X Interview interaction, $F(1, 45) = 4.046$, $p = .002$, η_p^2 .53. Post hoc tests revealed that Terrorist players reported being more deceptive in the Control interviews than in both the Strategic and Tactical interviews, $p = .004$, with no difference between the Strategic and Tactical interviews.

There was a significant main effect of player Veracity Group (builder; terrorist), $F(1, 145) = 44.847$, $p < .001$, η_p^2 .24, and Interview, $F(1, 145) = 11.010$, $p < .001$, η_p^2 .13 for difficulty. Terrorist players found the interviews more demanding than builder players. Post hoc tests revealed that participants found both the Tactical and strategic interviews more demanding than the Control interviews, $p < .009$, with no difference between Strategic and Tactical interviews, $p = .314$.

3.2 Police Observer Veracity Judgments

The initial post observation question asked observers to make a dichotomous choice as to whether the player was being deceptive or truthful (1 = *deceptive*; 2 = *truthful*). Percentage accuracy as a function of interview condition and veracity are displayed in Table II.

A series of 3 (interview) X 2 (participant veracity) repeated measures ANOVA were conducted on police observers' mean veracity performance (in each condition, for liars and truthteller observer participants were awarded a score ranging from 0 to 5 [i.e., six possibilities] according to the number of correct judgements made) and confidence scale data, applying Bonferroni corrections and followed by post hoc tests as appropriate.

For observer accuracy a significant main effect of Interview emerged, $F(2, 58) = 13.266$, $p < .001$, η_p^2 .89. Observer veracity (Lie; Truth) performance was significantly better in the TUE condition than both the SUE and Early (Control), $p < .001$, with no difference between the latter two conditions, $p = .975$. No interactions emerged, $ps > .018$ (see Table III).

A significant main effect of participant Veracity, $F(1, 29) = 46.412$, $p < .001$, η_p^2 .89, and of Interview, $F(2, 58) = 31.932$, $p < .001$, η_p^2 .87, emerged for the strength of judgment scale (1 = *definitely telling the truth* to 7 = *definitely lying*), The strength of judgments was significantly higher in the Early (Control) than in both the SUE and TUE conditions, all ps < .005 with no significant difference between the latter two conditions, $p = .40$. There was a non-significant interaction, $F = .828$, $p = .417$.

For confidence there was a significant main effect of Interview type (1 = *completely confident* to 7 = *not at all confident*), $F(2, 58) = 72.917$, $p < .001$, η_p^2 .91, and of participant Veracity, $F(1, 29) = 36.654$, $p < .001$, η_p^2 .98, and a significant interaction, $F(2, 58) = 21.168$, $p < .001$, η_p^2 .95. Overall, police observers were significantly more confident in their veracity judgments in the Early (Control) condition than both the SUE and TUE conditions all ps < .001 with no significant difference between the latter two conditions, $p = .313$ (see Table 4). Police observers were more confident when making lie judgments ($M = 4.81$, $SD = .80$) than truth judgments ($M = 2.81$, SD = .99), $p < .001$.

The strength of observers' confidence judgments of liars (Terrorists) were higher in the Early (Control) condition than in both the SUE and TUE all ps < .008 (see Table 4), with no difference between the latter two conditions. They were also more confident in their judgments of truth-tellers (Builders) in the Early (Control) condition than both the SUE and TUE conditions, all ps < .004, with no significant difference between the latter two conditions (see Table IV).

4. Discussion

We hypothesised that carefully managing the timing of evidence release would maximize opportunities for professional observers who had not been trained in either the SUE or TUE techniques to make more accurate veracity judgements. We argued that the incremental release of evidence as happens in TUE interviews should allow clearer separation of truthtellers and liars by observers because it allows truthtellers to evidence their innocence and highlights deceptive accounts more quickly. Our results support this hypothesis. Untrained police observers were significantly more accurate when making veracity judgments *per se* about interviewees in the TUE condition than in both the SUE and Early (Control) interviews. In the case of SUE and Early interviews overall performance was around chance (51% and 48%, respectively), whereas in the TUE condition performance was approaching 70%. Veracity performance for liars and truthtellers revealed a similar pattern of results. In the TUE condition, police observers were correct 67% of the time for liars and 70% for truthtellers. In both the SUE and Early (control) conditions performance hovered around chance for both liars and truthtellers, although veracity decisions for liars in SUE interviews was above chance (59%) albeit that this difference did not reach statistical significance.

These results lend further support to the psychological literature highlighting the importance of how and when to reveal evidence or any other relevant event information

during an investigative interview or formal conversation (Dando & Bull, 2011; Dando et al., 2015; Granhag et al., 2013; Luke et al., 2013; Luke, Hartwig, Joseph, et al., 2016; Ormerod & Dando, 2014; Sorochinski et al., 2014). As we have previously argued, it is equally as important for criminal justice that truthtellers are given the opportunity to evidence their innocence as quickly as possible as it is that deceivers are 'outed' by interviewing techniques. As well as supporting interviewers to better understand the veracity of accounts, it is also fundamental that professionals involved in criminal justice who have not been involved in the interview itself and who may not be expert interviewers themselves are assisted to make accurate veracity decisions.

One finding that on the face of it seemed counterintuitive was that police observers reported being far more confident about their decisions and rated their veracity judgements as being stronger following Early (control) interviews. Yet, they were significantly more accurate when making veracity decisions following TUE interviews. This pattern of results differs from previous findings where police interviewers had been trained to use the TUE technique. Here, they were, quite rightly, far more confident in their veracity decisions and rated their decisions more strongly, indicating that following training they recognized the value of the method for improving veracity performance.

Our data only allow us to speculate why the current officers generally lacked confidence in their performance. However, it seems sensible to assume that they were most comfortable and therefore most confident in their veracity decisions following Early (control) interviews because disclosing or revealing information/evidence at the start of an interview would have been more familiar to them. This is typical across many professional contexts whereby confidence and understanding of current practices can overshadow the efficacy of new methods (see Lewis, 2016). None of our police observers were specialist interviewers because we purposely sought to understand the efficacy of the techniques for non-specialist

professionals and so the more complex interviewing techniques taught to specialist interviewers such as SUE and TUE will have been unfamiliar.

TUE interviews follow a route map, which we believe was quickly and easily 'seen' and understood by police observers in this study, despite a lack of training. The piece by piece 'question-answer-reveal/challenge' OR 'question-answer -reveal/accept' pattern emerged quickly, and as we have previously argued this repetition is likely to support effective learning and decision-making in forensic contexts (Cacioppo & Petty, 1979; Dando & Bull, 2011; Kang, 2016; Nahari, 2018). This route map is not obvious in SUE interviews, which is an inherent weakness of the SUE technique for untrained professionals or layperson observers. It is unclear whether interviewees are being veridical until the very end of the interview when known information is revealed, by which time the concentration and cognitive resources of observers may have waned, or fast and furious decisions have been made which can be difficult to overturn (Dando & Ormerod, 2017; Walsh et al., 2018).

Again, as is the case in other domains such as eyewitness identification accuracy (Martschuk et al., 2019; Sauer et al., 2019) and witness memory performance (Goodwin et al., 2017; Loftus & Greenspan, 2017), confidence does not necessarily equate to competence and so confidence ratings of non-experts should be treated with caution even when collected immediately post interview. Not surprisingly, overall police observers were far more confident when making lie judgments than truth judgments and they rated lie judgments as being far stronger. Police experience deceptive communication more regularly than lay-persons and so typically are less likely to assume truthfulness. Consequently, police are usually more confident when making deception decisions than truth decisions (Levine et al., 1999; Masip et al., 2016; Masip & Herrero, 2017). However again, despite a lack of confidence when judging truthtellers in the TUE condition, they performed just as well as when judging liars.

This research is not without its limitations. We used a paradigm where participants created their own individual deceptions about an event they had actually taken part in. The event was complex and multifaceted, and as such mimicked the types of incidents that that police interviewers are concerned with on a day-to-day basis. Mock persons of interest (our participants) were not simply asked to maintain a deceptive statement presented by the experimenter. Rather, they had to think in real time about how to answer questions and how to behave to best prove their innocence or hide their deception.

That said, our mock persons of interest were not practiced deceivers, and they may have unknowingly revealed their status more readily than a professional criminal might. Equally, despite being motivated to carry out the experimenter instructions and being paid to do so, deceptive participants may have been less motivated to assert their innocence and/or hide their guilt than real persons of interest because of the lack of ramifications. Future researchers might consider extending our findings in a number of ways. First, by recruiting a cohort of practiced deceivers to play the game and then take part in the interview, perhaps from ex-offender populations, for example. Or, recruiting professionals who have day-to-day experience of working with deceivers to draw on their experiences during role play. Both cohorts of participants would support research to countenance some of the limitations of the current study.

Second, motivation could be enhanced by introducing a real-time financial reward and sacrifice system at various stages of the game and then throughout the interview process. For example, drawing on cooperative game theory, players could be incentivized to form coalitions (or teams) whereby the structure, strategies, and financial payoffs for the team as a whole relies on performance thereby introducing an additional motivational element ((Colman, 2003; Shubik, 2002). Nonetheless, our findings again highlight the utility of the

TUE technique for improving the detection of deception, but perhaps more importantly for

protecting the innocent.

References

Anolli, L., & Ciceri, R. (1997). The voice of deception;Vocal strategies of naive and able liars. *Journal of Nonverbal Behavior, 21*(4), 259-284.

Bull, R. (2014). When in interviews to disclose information to suspects and to challenge them? In R. Bull (Ed.) Investigative interviewing. New York: Springer.

Cacioppo, J. T., & Petty, R. E. (1979). Effects of message repetition and position on cognitive response, recall, and persuasion. *Journal of personality and Social Psychology, 37*(1), 97.

Colman, A. M. (2003). Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and Brain Sciences, 26*(2), 139-153.

Dando, C. J., & Bull, R. (2011). Maximising Opportunities to Detect Verbal Deception: Training Police Officers to Interview Tactically. *Journal of Investigative Psychology and Offender Profiling, 8*(2), 189-202. https://doi.org/10.1002/jip.145

Dando, C. J., Bull, R., Ormerod, T. C., & Sandham, A. L. (2015). Helping to sort the liars from the truth-tellers: The gradual revelation of information during investigative interviews. *Legal and Criminological Psychology, 20*(1), 114-128. https://doi.org/10.1111/lcrp.12016

Dando, C. J., & Ormerod, T. C. (2017). Analyzing decision logs to understand decision making in serious crime investigations. *Human factors, 59*(8), 1188-1203.

Dando, C. J., & Ormerod, T. C. (2019). Noncoercive human intelligence gathering. *Journal of Experimental Psychology: General*.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to Deception. *Psychological Bulletin, 129*(1), 74-118. https://doi.org/DOI: 10.1037/0033-2909.129.1.74

Goodwin, K. A., Hannah, P. J., Nicholl, M. C., & Ferri, J. M. (2017). The confident co☐ witness: The effects of misinformation on memory after collaborative discussion. *Applied Cognitive Psychology, 31*(2), 225-235.

Granhag, P. A., & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law, 14*(3), 189-200. https://doi.org/10.1080/10683160701645181

Granhag, P. A., Stromwall, L. A., Willen, R. M., & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology, 18*, 341-355.

Hartwig, M., Granhag, P. A., Strömwall, L., A. , & Kronkvist, O. (2006). Strategic Use of Evidence During Police Interviews: When Training to Detect Deception Works. *Law and Human Behavior, 30*(5), 603. http://proquest.umi.com/pqdlink?did=1195284781&Fmt=7&clientId=14829&RQT=309&VName=PQD

Hartwig, M., Granhag, P. A., & Stromwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology Crime & Law, 13*(2), 213-227. https://doi.org/10.1080/10683160600750264

Hartwig, M., Granhag, P. A., Stromwall, L. A., & Doering, N. (2010). Impression and Information Management: On the Strategic Self- Regulation of Innocent and Guilty Suspects. *The Open Criminology Journal, 3*(1), 10-16. https://doi.org/10.2174/1874917801003020010

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting Deception via Strategic Disclosure of Evidence. *Law and Human Behavior, 29*(4), 469-484. http://www.jstor.org/stable/4499433

Kang, S. H. (2016). Spaced repetition promotes efficient and effective learning: Policy implications for instruction. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 12-19.

Leins, D. A., Fisher, R. P., & Ross, S. J. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology, 18*(1), 141-151.

Levine, T. R. (2014). Active Deception Detection. *Policy Insights from the Behavioral and Brain Sciences, 1*(1), 122-128. https://doi.org/10.1177/2372732214548863

Levine, T. R., Park, H. S., & McCormack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the 'veracity effect'. *Communication Monographs, 66*, 125-144. https://doi.org/10.1080/03637759909376468

Loftus, E. F., & Greenspan, R. L. (2017). If I'm Certain, Is It True? Accuracy and Confidence in Eyewitness Memory. *Psychological Science in the Public Interest, 18*(1), 1-2. https://doi.org/10.1177/1529100617699241

Luke, T. J., Hartwig, M., Brimbal, L., Chan, G., Jordan, S., Joseph, E., Osborne, J., & Granhag, P. A. (2013). Interviewing to Elicit Cues to Deception: Improving Strategic Use of Evidence with General-To-Specific Framing of Evidence. *Journal of Police and Criminal Psychology, 28*(1), 54-62.

Luke, T. J., Hartwig, M., Joseph, E., Brimbal, L., Chan, G., Dawson, E., Jordan, S., Donovan, P., & Granhag, P. A. (2016). Training in the Strategic Use of Evidence technique: Improving deception detection accuracy of American law enforcement officers. *Journal of Police and Criminal Psychology, 31*(4), 270-278.

Luke, T. J., Hartwig, M., Shamash, B., & Granhag, P. A. (2016). Countermeasures Against the Strategic Use of Evidence Technique; Effects on Suspects Strategies. *Journal of Investigative Psychology & Offender Profiling, 13*(2), 131-147. https://doi.org/10.1002/jip.1448

Martschuk, N., Sporer, S. L., & Sauerland, M. (2019). Confidence of Older Eyewitnesses: Is It Diagnostic of Identification Accuracy? *Open Psychology, 1*(1), 132-151.

Masip, J., Alonso, H., Herrero, C., & Garrido, E. (2016). Experienced and novice officers' generalized communication suspicion and veracity judgments. *Law and Human Behavior, 40*(2), 169.

Masip, J., & Herrero, C. (2017). Examining police officers' response bias in judging veracity. *Psicothema, 29*(4), 490-495.

Nahari, G. (2018). The applicability of the verifiability approach to the real world. In *Detecting concealed information and deception* (pp. 329-349). Elsevier.

Ormerod, T. C., & Dando, C. J. (2014). Finding a needle in a haystack: Towards a psychologically informed method for aviation security screening. *Journal of Experimental Psychology: General, 144*(1).

Parkhouse, T., & Ormerod, T. C. (2018). Unanticipated questions can yield unanticipated outcomes in investigative interviews. *Plos One*, 1-22. https://doi.org/10.1371/journal.pone.0208751

Sandham, A. L., Ormerod, T. C., Dando, C. J., Bull, R., Jackson, M., & Goulding, J. (2011). Scent Trails: Countering Terrorism through Informed Surveillance *Lecture Notes in Computer Science* Engineering Psychology and Cognitive Ergonomics, 14th HCI International, July 2011, Orlando, Florida.

Sandham, A. L., Ormerod, T. C., Dando, C. J., & Menacere, T. (2015). On the Trail of the Terrorist: A Research Environment to Simulate Criminal Investigations. In A. L. Stedman, G. (Ed.), *Hostile Intent and Counter☐Terrorism - Human Factors Theory and Application* (pp. 129 - 146). Ashgate Publishing. https://doi.org/10.1201/9781315587080-11

Sauer, J. D., Palmer, M. A., & Brewer, N. (2019). Pitfalls in using eyewitness confidence to diagnose the accuracy of an individual identification decision. *Psychology, Public Policy, and Law, 25*(3), 147-165. https://doi.org/10.1037/law0000203

Shubik, M. (2002). The uses of teaching games in game theory classes and some experimental games. *Simulation & Gaming, 33*(2), 139-156.

Sorochinski, M., Hartwig, M., Osborne, J., Wilkins, E., Marsh, J., Kazakov, D., & Granhag, P. A. (2014). Interviewing to detect deception: When to disclose the evidence? *Journal of Police and Criminal Psychology, 29*(2), 87-94.

Sporer, S. L. (2016). Deception and Cognitive Load: Expanding Our Horizons with a Working Memory Model. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00420

Verigin, B. L., Meijer, E. H., Bogaard, G., & Vrij, A. (2019). Lie prevalence, lie characteristics and strategies of sefl reported good liars. *Plos One, 14*(12). https://doi.org/10.1371/journal.pone.0225566

Vrij, A. (2019). Deception and truth detection when analyzing nonverbal and verbal cues. *Applied Cognitive Psychology, 33*(2), 160-167.

Walsh, D., Dando, C. J., & Ormerod, T. C. (2018). Triage decision-making by welfare fraud investigators. *Journal of applied research in memory and cognition, 7*(1), 82-91.

Table I. *Mock suspect post interview perceptions (N = 151)*

Interview Condition	Motivation M (SD)	Deceptiveness M (SD)	Difficulty M (SD)
TUE	5.33 (1.64)	2.76 (1.75)	4.31 (.754)
Truth-teller (Builder)	5.23 (1.55)	1.67 (1.01)	3.44 (1.22)
Liar (Terrorist)	5.49 (1.33)	3.99 (2.10)	5.43 (1.01)
SUE	5.49 (1.28)	3.23 (2.09)	3.90 (1.08)
Truth-teller (Builder)	5.61 (1.19)	1.32 (.99)	3.42 (1.12)
Liar (Terrorist)	5.70 (1.48)	4.84 (2.01)	4.68 (.89)
Early (Control)	5.48 (1.39)	3.38 (1.45)	3.08 (1.14)
Truth-teller (Builder)	5.68 (1.29)	1.77 (.79)	2.58 (.99)
Liar (Terrorist)	5.19 (1.33)	5.10 (1.98)	3.81 (1.20)

Table II. *Percentage accuracy as a function of interview and veracity*

% Accuracy	Interview Condition		
	TUE	SUE	Early (Control)
Liar (Terrorist)	67	59	50
Truth-teller (Builder)	70	41	45
Overall	68	51	48

Table III. *Untrained police observers mean veracity performance as a function of interview condition.*

Interview Condition	Liar *M* (95% CI)	Truth-Teller *M* (95% CI)
TUE	3.20 (2.84; 3.56)	3.50 (2.81; 3.79)
SUE	2.20 (1.85; 2.56)	2.67 (2.28; 2.05)
Early (Control)	2.40 (2.04; 2.76)	2.53 (2.98; 2.97)

Table IV. *Untrained police observers' mean strength of judgment and confidence ratings.*

Interview Condition	Strength *M* (SD)	Confidence *M* (SD)
TUE	4.66 (.82)	4.53 (.98)
Truth-teller (Builder)	5.13 (.49)	5.13 (.49)
Liar (Terrorist)	4.93 (.79)	4.81 (1.11)
SUE	4.73 (.71)	4.47 (1.08)
Truth-teller (Builder)	4.36 (.57)	4.36 (.57)
Liar (Terrorist)	4.06 (1.22)	3.29 (.59)
Early (Control)	3.78 (.89)	3.40 (1.01)
Truth-teller (Builder)	3.51 (.63)	3.51 (.63)
Liar (Terrorist)	4.27 (.84)	3.79 (.64)