Journal of Investigative Psychology and Offender Profiling
J. Investig. Psych. Offender Profil. (2015)
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI**: 10.1002/jip.1443

Train the Trainers: A First Step towards a Science-Based Cognitive Lie Detection Training Workshop Delivered by a Practitioner

ALDERT VRIJ*, SAMANTHA MANN, SHARON LEAL, ZARAH VERNHAM and MARTIN VAUGHAN

Psychology Department, University of Portsmouth, Portsmouth, UK

Abstract

A training workshop utilising recent research in cognitive lie detection was designed and evaluated. A unique future of the training was that a practitioner (retired police detective) instead of scientists (e.g. the training developers) introduced the techniques. To evaluate the training, 27 experienced police detectives each interviewed one mock suspect (a truth teller or liar) before training, and another group of 23 experienced police detectives interviewed one mock suspect (a truth teller or liar) after training. The police detectives were free to interview the mock suspect in any way they felt appropriate, but those who had received training were asked to try to incorporate (some of) the taught techniques in their interviews. The detectives made veracity judgements, and the interviews were transcribed and coded for the amount of detail elicited and the questions asked. Training had a modest effect on the ability to distinguish between truths and lies but resulted in a higher percentage of appropriate questions asked. Trainees did not implement the taught techniques to an equal extent, but when they did, the techniques enhanced the elicitation of information. The training study also revealed challenges, particularly difficulty in implementing the taught techniques into practice and asking the right questions to elicit differences in detail between truth tellers and liars. Copyright © 2015 John Wiley & Sons, Ltd.

Key words: lie detection; training; interviewing to detect deception

Meta-analyses of the deception literature revealed that (non-)verbal cues to deceit are typically faint and unreliable (DePaulo *et al*., 2003) and that observers who rely on non-verbal cues or on verbal cues without using established verbal lie detection procedures such as criteria-based content analysis (Steller & Köhnken, 1989; Vrij, 2005) or reality monitoring (Masip, Sporer, Garrido, & Herrero, 2005; Sporer, 2004) barely exceed the level of chance in correctly classifying truth tellers and liars (Bond & DePaulo, 2006). Based on these meta-analyses, some researchers started to examine whether investigators can *elicit new* cues or *enhance existing* cues to deceit through specific interview protocols

*Correspondence to: Aldert Vrij, Psychology Department, University of Portsmouth, King Henry Building, King Henry 1 Street, Portsmouth, Hants, PO1 2DY, UK.
E-mail: aldert.vrij@port.ac.uk

(Vrij & Granhag, 2012). The two most extensively examined approaches to date are the strategic use of evidence technique (Granhag & Hartwig, 2015; Hartwig, Granhag, & Luke, 2014) and the cognitive lie detection approach (Vrij, 2015).

We developed a cognitive lie detection training workshop that took into account the results of more than 20 studies into cognitive lie detection (Vrij, 2015). In a first training evaluation study, the developers (all scientists) delivered the training to experienced police detectives (Vrij, Leal, Mann, Vernham, & Brankaert, 2015). Training improved the ability to (1) detect deceit and (2) elicit information. Also, some challenges arose, particularly facing difficulty in implementing the taught techniques into practice. In the present experiment, the same training was delivered and evaluated, but the techniques were introduced by a retired police detective rather than by scientists. It is important to know whether non-scientists can deliver a science-based training, because it gives insight into the question of whether training can be delivered on a wider scale.

The training was relatively extensive compared with other deception detection training workshops (see Hauch, Sporer, Michael, & Meissner, 2014, for a meta-analysis of such training studies). It lasted for 7 hours, and none of the 30 studies included in Hauch *et al*. (2014) lasted so long. In addition, the interview techniques, which were introduced in the training, exclusively focussed on verbal cues to deceit. A training based on verbal cues typically yields the best results (Hauch *et al*., 2014). In terms of training evaluation, police detectives interviewed mock suspects without being trained or after training, and the detectives were invited to interview the mock suspects in any way they felt appropriate. This is a more realistic evaluation of a training than asking untrained and trained observers to passively watch videotapes of truth tellers and liars and ask them to make veracity judgements (the method used in the vast majority of training studies).

THE COGNITIVE LIE DETECTION APPROACH

The core of the cognitive lie detection approach is that investigators can *magnify* the differences in cognitive cues displayed by truth tellers and liars through interventions that make the liar's task more cognitively demanding. The cognitive lie detection approach comprises three components: (1) imposing cognitive load; (2) encouraging interviewees to say more; and (3) asking unexpected questions.

Imposing cognitive load is based on the empirical finding that in *interview settings* lying is typically more mentally taxing than truth telling (see for example functional magnetic resonance imaging research [Christ, Van Essen, Watson, Brubaker, & McDermott, 2009; Farah, Hutchinson, Phelps, & Wagner, 2014; Lisofsky, Kazzer, Heekeren, & Prehn, 2014; Vrij & Ganis, 2014], as well as other sources of evidence, including analyses of behaviour and self reports [Vrij, Fisher, Mann, & Leal, 2006]). Imposing cognitive load refers to investigators' interventions aimed at making the interview setting mentally more difficult. Liars, who require more cognitive resources than truth tellers, will have fewer cognitive resources left over. If cognitive demand is further raised, liars may be less able than truth tellers to cope with these additional requests (Vrij, Granhag, Mann, & Leal, 2011). One way to impose cognitive load is by asking interviewees to tell their stories in reverse order. This results in noticeable cues to deceit and facilitates lie detection (e.g. Evans, Meissner, Michael, & Brandon, 2013; Vrij *et al*., 2008). Another example is 'forced turn-taking', a technique that can be applied when more than one person is interviewed at the same time (groups of liars and groups of truth tellers). In the forced turn-taking

DOI: 10.1002/jip

technique, the interviewer asks one interviewee to start answering a question. The interviewer then interrupts that person after a short period of time and asks a second interviewee to continue with the story. The interviewer then also interrupts this person after a short period of time and asks a third interviewee (or the first interviewee again in case of interviewing a pair) to continue, and so on. Forced turn-taking leads to noticeable cues to deceit. Truth-telling pairs continued on from one another, whereas lying pairs waited and repeated previously said information before continuing (Vernham, Vrij, Mann, Leal, & Hillman, 2014). Forced turn-taking also facilitates distinguishing between truth tellers and liars (Vernham *et al*., 2014). A third way of imposing cognitive load is by engaging interviewees in a second task (e.g. watching a video of an unrelated event) whilst conducting the interview (Johnston, Greenberg, Fisher, & Martin, 1970). Doing two tasks simultaneously (watching a video and telling a story) is cognitively more demanding than doing one task (i.e. just storytelling), and liars find it more difficult to cope with the addition of such a second task than truth tellers (Debey, Verschuere, & Crombez, 2012; Visu-Petra, Varga, Miclea, & Visu-Petra, 2013).

The second cluster of techniques aims to encourage interviewees to provide more information. Such techniques are in the interest of truth tellers. If truth tellers provide much information, they are more likely to be believed, because the richer an account is perceived to be in detail, the more likely it is to be believed (Bell & Loftus, 1989; Johnson, 2006). Moreover, the additional information truth tellers provide could give leads for investigators to check. There are reasons to assume that techniques to encourage interviewees to say more have a lesser effect on liars, because liars have some constraints. They may be reluctant to be detailed as they run the risk that such detail can be falsified by an investigator (Hartwig, Granhag, & Strömwall, 2007; Nahari, Vrij, & Fisher, 2014) and they may lack the imagination to conjure up details that sound plausible (Leal, Vrij, Warmelink, Vernham, & Fisher, 2015). Liars may also want to limit the amount of false information they provide so that they have less false information to remember in case they are interviewed again (Vrij, 2008). Hence, it can be expected that techniques to encourage interviewees to say more lead to truth tellers adding more (plausible) detail than liars. Empirical research has supported this premise, and methods to encourage interviewees to say more discussed in the training include the use of (1) a model statement (playing the participant a very detailed statement as an example) (Leal *et al*., 2015); (2) a supportive interviewer (e.g. Mann *et al*., 2013; Shaw, Vrij, Mann, Leal, & Hillman, 2013); (3) deliberate mimicry of the interviewee (Shaw *et al*., 2015); and (4) drawings (e.g. Roos af Hjelmsäter, Öhman, Granhag, & Vrij, 2014; Vrij *et al*., (2012)).

The third cluster of techniques relates to asking unexpected questions. Liars prepare themselves for anticipated interviews by preparing answers to questions they expect to be asked (e.g. Hartwig *et al*., 2007). This strategy makes sense. Planning makes lying easier, and planned lies typically contain fewer cues to deceit than spontaneous lies (DePaulo *et al*., 2003). This strategy, however, will be fruitful only if liars correctly anticipate which questions will be asked. Investigators can exploit this limitation by asking questions that liars do not anticipate. Although liars can refuse to answer unexpected questions by saying 'I don't know' or 'I can't remember', such responses will create suspicion if they are about central aspects of the target event. A liar, therefore, has little option other than to fabricate a plausible answer on the spot, which is cognitively demanding.

For liars, expected questions should be easier to answer than unexpected questions, because liars can give their planned and rehearsed answers to the expected questions but they need to fabricate answers to the unexpected questions. This may result in answers

DOI: 10.1002/jip

to the unexpected questions that are less detailed and less plausible than the answers to the expected questions. In contrast, truth tellers experience similar levels of cognitive load whilst answering expected and unexpected questions, and they should produce more comparable answers with the expected and unexpected questions than liars. Research supports this premise, and examples of unexpected questions include spatial questions (Vrij *et al*., 2009), questions about processes (e.g. planning of a trip) rather than outcomes (e.g. purpose of a trip) (Mac Giolla, Granhag, & Liu-Jönsson, 2013), and asking of the same question twice in different formats (Leins, Fisher, & Vrij, 2012; Leins, Fisher, Vrij, Leal, & Mann, 2011).

TRAINING PERFORMANCE INDICATORS

To evaluate the training, we examined three performance indicators: (1) accuracy in distinguishing truth tellers from liars; (2) the total amount of detail provided by the mock suspects; and (3) the types of questions the trainees (police detectives) asked during the interviews.

In lie detection studies, observers make veracity judgements about target persons they observe. The target person is either telling the truth or lying, so the observers can expect a 50% accuracy if they just guess. Bond and DePaulo (2006) reviewed the available lie detection research in their meta-analysis and reported an average accuracy rate of 54% in correctly classifying truth tellers and liars. Vrij (2008) examined the accuracy rates obtained by professionals (e.g. police officers, police detectives, customs officers, and secret service agents) in 30 lie detection studies. The average accuracy rate was 56% for detecting truths and 56% for detecting lies (56% total accuracy). In 29 of those studies, observers passively watched video fragments of truth tellers and liars. This deviates from what practitioners typically do in real life, as they often actually interview people. In the remaining study, police detectives actually interviewed (mock) suspects and achieved a 57% accuracy rate (Hartwig, Granhag, Strömwall, & Kronkvist, 2006). We thus can anticipate a total accuracy rate of around 56% amongst untrained interviewers but expect a higher percentage amongst trained interviewers (Hypothesis 1).

Deception research has shown that truth tellers typically give more detail than liars (DePaulo *et al*., 2003; Vrij, 2008). Liars may lack the imagination to conjure up details that sound plausible, may be reluctant to give detail as they run the risk that such detail can be proven false by an investigator, and may want to limit the amount of false information they provide so that they have less false information to remember and report in case they are interviewed again (Vrij, 2008). We thus predicted that truth tellers will provide more detail than liars, but because the training is meant to elicit or enhance differences between truth tellers and liars, we predicted that this would particularly be the case in interviews with the trained interviewers (Hypothesis 2).

The third performance indicator was the type of questions asked. For this purpose, we used the distinction between *appropriate* and *inappropriate* questions as introduced by Griffiths and Milne (2006) and used by Oxburgh, Ost, and Cherryman (2012). Appropriate questions are open-ended questions, probing questions, and encourager/acknowledgement questions, whereas inappropriate questions are closed questions, forced-choice questions, leading questions, multiple questions at once, opinion/statement questions, and echo questions. Appropriate questions are the most productive in terms of eliciting information (Fisher, Falkner, Trevisan, & McCauley, 2000; Griffiths & Milne, 2006; Oxburgh *et al*., 2012), because they elicit free recall, allow interviewees to collect their thoughts in their

DOI: 10.1002/jip

own way, and give interviewees time to think (Powell, Fisher, & Wright, 2005; Snook, Luther, Quinlan, & Milne, 2012).

Inappropriate questions lead to less information (Myklebust & Bjorklund, 2006; Oxburgh *et al*., 2012) and are considered inappropriate because they are suggestive. In addition, the often misleading information embedded in these questions can be incorporated into a person's memory and could eventually lead to false recall in later stages of the interview process (Griffiths & Milne, 2006; Gudjonsson & Clark, 1986).

Oxburgh *et al*. (2012) found that only 29% of the questions asked were appropriate, which reflects a common trend found in research (Fisher, 2010; Vrij, Hope, & Fisher, 2014; Walsh & Bull, 2012). To explain the relatively infrequent use of appropriate questions in investigative interviewing, Oxburgh and Dando (2011) gave the following five explanations: (1) police training manuals differ in their definitions of open-ended and closed questions, which could cause confusion; (2) an interviewer is inclined to maintain control over the interview, which could be established by asking closed questions; (3) if in a specific interview an interviewer mainly seeks confirmation of known facts, asking closed questions is a fast way of achieving that information; (4) interviewers may seek power over the interviewee, and asking closed questions is a way to achieve this; and (5) asking open-ended questions is an unfamiliar way of questioning as in everyday interactions people tend to use closed rather than open-ended questions.

Based on the findings of Oxburgh *et al*. in England and Wales, we expected that around 29% of questions asked would be appropriate. The techniques employed in the cognitive lie detection approach solely consist of appropriate questions (mainly open-ended questions), and as such, we expected trained detectives to use more appropriate questions than untrained detectives, and, in particular, more open-ended questions (Hypothesis 3). Four of the taught techniques can easily be detected by reading the transcripts: (1) model statement; (2) reverse order; (3) drawings; and (4) spatial questions. We examined whether using these four techniques in the interviews would be related to eliciting information and lie detection. Because these techniques have been shown to elicit information in interviews and differentiate truth tellers from liars, we expected that this would be the case and therefore predicted that questions related to the taught techniques would elicit more information than the other questions (Hypothesis 4a) and would facilitate lie detection (Hypothesis 4b).

METHOD

Participants

Interviewers

A total of 50 police detectives (30 men) took part. Their age ranged from 27 to 58 years with an average age of $M = 43.62$ ($SD = 9.50$), and their length of service ranged from 7 to 39 years with an average length of $M = 20.38$ ($SD = 7.30$). The control ($n = 27$) and experimental ($n = 23$) groups did not differ from each other on these three background characteristics (all p's $> .101$). In England and Wales, there are five tiers of interview training for officers (Tier 5 is the highest level of training). Most detectives ($n = 23$) were Tier 3 trained for interviewing suspects, 13 interviewers were Tier 2 trained, one officer was Tier 4 trained, and 11 officers were Tier 5 trained (two officers did not indicate their training level). This level of training had no effect on the accuracy rates reported in the Results section or on the type of questions asked. Thirty-nine detectives judged themselves as

DOI: 10.1002/jip

experienced in interviewing suspects (a score of 4 or higher on a 5-point Likert scale). The control and experimental groups did not differ from each other on those two experience characteristics either (both p's $> .928$).

Interviewees

A total of 50 interviewees (35 women) acted as mock suspects in the training study of which 27 took part in the control group and 23 in the experimental group. The sample was made up of undergraduate students and university staff. Their age ranged from 18 to 61 years with an average age of $M = 31.80$ ($SD = 13.85$). The control and experimental groups did not differ from each other on those two characteristics (both p's $> .259$).

THE TRAINING WORKSHOP

Police detectives were recruited through police detective training coordinator associates. The training content was briefly explained in an email that was sent to approximately 100 police detectives through these coordinators with the request to email the first author directly if interested in attending. The training and study was held over 5 days, and on each day, between four and six trainees participated.

The training comprised two elements. In the initial training, the training developers (A. V., S. M., and S. L., all authors on this paper) trained the trainees. That is, A. V. was responsible for presenting the background information via a PowerPoint presentation, and S. M. and S. L. were responsible for carrying out the exercises and eliciting discussions amongst trainees. For the 'train-the-trainer' element, we trained in a 1-day session a retired Welsh police detective to present the PowerPoint presentation, thus, in effect, to take over A. V.'s role in the training. This trainer had previously attended and hence experienced the course as a trainee. In the 'train-the-trainer' part, the retired police detective gave the background information using our PowerPoint slides. A. V., S. M., and S. L. were present during the training and assisted the new trainer if trainees asked questions. S. M. and S. L. continued carrying out the exercises and elicited discussions amongst trainees.

The training started at 9 AM with a 1-hour discussion of the 'pitfalls in lie detection' section derived from Vrij (2008) and Vrij, Granhag, and Porter (2010). The pitfalls section consists of difficulties in lie detection (e.g. a cue akin to Pinocchio's growing nose does not exist, liars try to fool investigators through countermeasures, liars embed their lies in truthful stories, lack of adequate feedback about the accuracy of veracity judgements made in real life, and some people are good liars), followed by errors commonly made by investigators (e.g. the incorrect use of heuristics, paying attention to the wrong cues [e.g. increase of movements whereas liars typically move less than truth tellers], the Othello error—interpretation of signs of nervousness in truth tellers as a sign of deceit, overemphasis on non-verbal cues, and failing to take into account inter-individual and intra-individual differences).

The pitfalls section was followed by a 6-hour 'Opportunities in lie detection' section (the cognitive lie detection approach), which included the introduction of 10 techniques as well as demonstrations, videos, and exercises. Table 1 provides brief descriptions of the 10 techniques taught in the training, whereas Appendix 1 provides brief descriptions of the demonstrations, videos, and exercises. Several short coffee breaks and a half-hour lunch break were included in this 'Opportunities in lie detection' section.

DOI: 10.1002/jip

Table 1. Interviewers' evaluation of the taught techniques

	Used in study (*N*)
Imposing cognitive load	
Reverse order (Recall an event in reverse chronological order.)	11
Secondary task (Carry out two tasks simultaneously: storytelling and a second independent task.)	5
Forced turn-taking (One interviewee starts answering the question, and after a short period of time, a second interviewee is asked to continue with the story. After a short period of time, a third person (or the first person again) is asked to continue the story, and so on.)	0
Encouraging interviewees to say more	
Model statement (Let the interviewee listen to a very detailed answer to a question unrelated to the event under investigation.)	13
Supportive interviewer (An interviewer who smiles and nods his or her head.)	5
Mimicry (The interviewer mimics the posture/seating position of the interviewee.)	4
Drawings (Request to sketch—for example a location or an object—rather than providing a verbal recall of that location or object.)	22
Asking unexpected questions	
Spatial questions (Questions about the spatial arrangement of people and objects: 'Where did you and your friend sit in the restaurant?' and 'Standing in the doorway looking to your left. What you could see?')	18
Processes versus outcomes (Outcomes refer to the end result, 'What are you going to do tonight?' and 'What is the purpose of your trip?', whereas processes refer to reaching this end result, 'What made you decide to go to this particular film tonight?' and 'What did you do to plan this trip?')	4
Asking the same question twice in a different format (Ask someone to (1) verbally recall the layout of a location followed by the request to (2) sketch the layout of that location.)	7

TRAINING EVALUATION

Procedure

Interviewers

On arrival at 8.30 AM, the police detectives were invited to complete a brief questionnaire about themselves (age, gender, and experience in interviewing). Officers in the *control condition* were then asked to interview a mock suspect in the way they would normally do (or as close as they could to normal given the experimental procedure). They were told that they had 30 minutes maximum for the interview and that the interview would be audiotaped. They were given background information about the mock-crime scenario (detailed later). After the interview, they completed a short questionnaire in which they made a dichotomous veracity judgement ('What do you think the veracity status of the suspect was? Truth teller or liar'). They then attended the training, which was delivered by the training developers (A. V., S. L., and S. M.), and started at 10 AM.

Officers in the experimental condition first attended the training delivered by the retired police detective. After the training (which finished at 4 PM), they were asked to interview a mock suspect and were asked, if possible, to incorporate (some of) the techniques taught in the training. Similar to the detectives in the control condition, they were told that they had 30 minutes to interview the mock suspect and that the interview would be audiotaped. They were given the same background information about the mock-crime scenario as the

DOI: 10.1002/jip

detectives in the control condition and completed the same post-interview questionnaire. In addition, we also asked the detectives whether they had used the techniques taught in the training (yes/no) and gave them hereby a list of the 10 techniques taught. Four different interview rooms were available for the police detectives, which means that four interviews could take place at the same time. The detectives were randomly allocated to the mock suspects.

We used two rather than one mock-crime scenarios in the study (the restaurant scenario and secret meeting scenario) to decrease the likelihood that any outcomes of the evaluation could be scenario specific. The restaurant scenario was derived from Strömwall, Granhag, and Jonsson (2003) and Vrij *et al*. (2009) but differed from those studies in that participants stayed in the restaurant individually rather than in pairs. The secret meeting scenario is derived from Shaw, Vrij, Leal, & Mann (2014) but differed from that study in that the participant saw a videotape of the meeting rather than attended the meeting.

The following information was given about the scenarios to the police detectives:

Restaurant scenario. A blue computer tablet has been stolen from an office in the last 15 minutes, and your role is to examine this crime. The suspect you are going to interview has either spent the last 15 minutes in a local restaurant called (name restaurant) (truth tellers) or went briefly to that restaurant, stole the computer tablet after leaving that restaurant, and pretended to have stayed in the restaurant for the last 15 minutes (liars). The participant went to the restaurant with a person called Gary to meet another person, Chris, who was not actually there. Your task is twofold: (1) try to obtain as much information as possible from the suspect about his or her stay at the restaurant and (2) decipher whether she or he has taken the computer tablet from an office in this building. Remember, liars also have visited the restaurant, albeit briefly. We will now give you pictures of (1) the interior of the restaurant where the suspect has been (truth tellers for about 15 minutes and liars only briefly) and (2) the office from where the computer tablet was taken. Truth tellers are aware that they are suspected of having stolen a computer tablet from this building.

Secret meeting scenario. You are a member of an intelligence agency who is going to interview an individual who has seen video footage of a secret meeting that is of great importance to the agency but that is now lost. In that meeting, a possible location is discussed where to install a spy device. Your task is to obtain as much information as possible about (1) the spy device and (2) where it will be installed. It could be that the interviewee has been told that she or he can trust you, and if so, the interviewee will give you accurate information. It could also be that the interviewee has been told that she or he cannot trust you, and if so, the interviewee will give you a mixture of accurate and inaccurate information. Although you have no information about the location and device, the interviewee is unaware of this and believes that you have some information, but she or he has no idea what this information actually is. She or he has therefore been told that the best strategy is to give you a mixture of accurate and inaccurate information. Your task is twofold: (1) try to obtain as much information about the meeting from the interviewee as possible and (2) decipher whether she or he is lying to you. Remember, liars will not lie to you all the time but will give you a mixture of accurate and inaccurate information.

Interviewees

Participants were recruited by email and online advertisements at the university. Interviewees were unaware that the interviewers were taking part in a cognitive lie detection training workshop, neither did they know that the interviewers were experienced police

DOI: 10.1002/jip

detectives. On arrival, participants were randomly allocated to the restaurant or secret meeting scenario and to the truth or lie condition. A confederate took participants in the *restaurant scenario* to a certain restaurant under the ruse of meeting another person. When arriving in the restaurant, the confederate told the participant that the other person was not there. The confederate bought a drink for the participant and himself and asked the participant to sit down and to wait for him to return (truth tellers) or went straight back with the participant to the experimenter room, where the experimenter instructed the participant to steal a computer tablet from an office (liars). After 15 minutes, the confederate collected the participant (truth teller) from the restaurant and brought the participant to the experimenter room.

Truth tellers were then told a computer tablet was stolen in the building and that they were one of the suspects. They would be interviewed about their whereabouts in the last 15 minutes, the time they were in the restaurant, and were asked to tell the interviewer that they were in the restaurant and to convince the interviewer that they were there.

Liars were also told that they were suspected of having taken the computer tablet and that they would be interviewed about their whereabouts during the last 15 minutes. They were instructed to deny having taken the computer tablet but to tell the interviewer that they were in (name restaurant) during the last 15 minutes (the restaurant they briefly went to with the experimenter). The reason for staying in the restaurant was that the experimenter asked them to stay there and to wait for him.

Participants in the *secret meeting scenario* took on the role of a security officer and saw video footage of a secret meeting between three people in which a vote was taken on a suitable location to plant a spy device. In the video, two locations were discussed, and a third location was mentioned but not discussed because one person had to leave. Once the video had finished, the participants were allocated to the truth-telling or lying condition. *Truth tellers* were informed that the footage they had just watched had disappeared and the agency had launched an investigation. The agency believed they had a mole working for them and it was of the upmost importance that the investigators knew as much detail about the video as they could. Truth tellers were told to fully cooperate with the investigators, to be completely truthful, and to answer the questions to the best of their knowledge.

Liars were informed that the footage they had just viewed had disappeared and that the agency had launched an investigation because they needed to know in as much detail as possible what had happened in the video. Liars were told it was now their responsibility to recall that information in an interview and that the intelligence agency believes they have a mole working for them. This mole could be the investigators the liars are going to talk to, so the information cannot be disclosed to them. Liars were told the investigators knew the device would be placed somewhere, but they do not know where. So, above all, they must *not* reveal the location that was selected to hide the spy device, and their objective was to mislead the investigators. Liars were told, when asked to describe the location that was *selected*, they must provide some false, decoy information. They were told to use the third location, which was not discussed, as the location that was selected to plant the device. As no information was provided in the meeting about a third location, liars needed to invent these details. In total, they needed to make up three bits of information: first, the location of the building where the device would be planted. Second, within that building specifically where the device would be planted, and thirdly, a reason why it is a suitable location. Liars were also told they needed to mislead the investigators about the device. The investigators know something about the device, but they do not have all the details, and it is not clear what they know. Because of this, liars needed to provide *some truthful*

DOI: 10.1002/jip

and *some false* information about the device. This will help them appear cooperative without having to tell them everything. How much truthful and false information they gave was up to the participants.

Truth tellers and liars in both scenarios were then told that it is important to convince the interviewer that they were truthful and that they would receive £10 as a reward if they did so. If the participant did not convince the interviewer, she or he would have to write a report about his or her whereabouts during the last 15 minutes. The participant was then taken to the interview room where the interview took place.

After the interview, participants completed a post-interview questionnaire, which measured motivation, likelihood of receiving £10, likelihood of writing a statement, and how difficult they thought the interview was. To measure motivation, participants were asked to what extent they were motivated to perform well in the interview on a 5-point Likert scale (1 = *not at all motivated* to 5 = *very motivated*). Likelihood of receiving the £10 or writing a statement was measured on 7-point Likert scales (1 = *not at all* to 7 = *totally*). Cognitive load experienced during the interview was measured with three questions: (1) 'I felt that the interview required a lot of thinking'; (2) 'I felt that the interview was mentally difficult'; and (3) 'I had to concentrate a lot during the interview'. Answers were given on 7-point Likert scales (1 = *disagree* to 7 = *agree*). These three questions were clustered into one 'cognitive load' index (Cronbach's alpha = .693).

The interviews in both scenarios were audio recorded and subsequently transcribed. All interviewers and interviewees signed informed consent forms prior to the study and were given a debriefing form after the study. All interviewees were given £10.

Coding

All coders were blind to the hypotheses and experimental conditions of the study.

Total detail

A coder read the transcripts and coded them for number of details. Detail included all the perceptual details (information about what the examinee saw or heard), spatial details (information about the spatial arrangement of people and/or objects), and temporal details (information about when the event happened or an explicit description of a sequence of events). We did not split detail into these sub-categories as no hypotheses were formulated about them. Thus the sentence 'There was a black napkin on the table' would be coded as four details. A second coder coded a random sample of 10 transcripts (19%). Inter-rater reliability between the two coders was excellent (single-measures intraclass correlation coefficient [ICC] = 0.95).

Questions asked

For coding the questions asked by the detectives in the interviews, we used Griffiths and Milne's (2006) categorisation of question types in forensic investigative interviewing: (1) open-ended questions (questions mostly beginning with 'Tell', 'Describe', or 'Explain'); (2) probing questions (questions beginning with one of the five WH words 'who', 'where', 'when', 'which', or 'why' or beginning with 'how'); (3) closed questions (questions that can only be answered with 'yes' or 'no'); (4) leading questions (where the question suggests an answer to the interviewee); (5) forced-choice questions (questions that force the interviewee to choose between two or more options); (6) echo questions (repeating the

DOI: 10.1002/jip

information given by the interviewee but phrased as a question); (7) encourager or acknowledge questions (questions or statements that encourage interviewees to continue talking like 'Mmmm', 'Uh-huh', 'Ah right, thank you', and 'Yeah, carry on'); (8) multiple questions at once (multiple questions at the same time without giving the interviewee the opportunity to respond between the questions); and (9) opinion or statement questions (expressions of opinions or statements put to the interviewee) (Snook *et al*., 2012).

Following Oxburgh *et al*. (2012), a coder read the transcripts and coded them for the questions asked and made a distinction between *appropriate questions* (open-ended questions, probing questions, and encourager/acknowledgement questions) and *inappropriate questions* (closed questions, forced-choice questions, leading questions, multiple questions at once, opinion/statement questions, and echo questions). A second coder coded a random sample of 15 transcripts. Inter-rater reliability between the two coders was excellent (single-measures ICC = 0.99 for both appropriate and inappropriate questions).

For four techniques, which can be easily spotted by reading the transcripts ([1] model statement; [2] reverse order; [3] drawings; and [4] spatial questions), one of the trainers coded the number of times the detectives had used each of those techniques in the way we taught them to use it. For example, for the use of the model statement, the detective first needed to elicit a free recall from the participant via an open-ended question (followed by further questioning, optional), followed by the model statement and then again an open-ended question eliciting a free recall. For the use of drawings, the detective should have asked the mock suspect 'Please sketch everything you could see when…' (Vrij, Leal *et al*., 2010). A second coder coded a random sample of 10 transcripts. Inter-rater reliability between the two coders for the cluster of four techniques was excellent (single-measures ICC = 0.90). The number of questions based on these four techniques asked in the post-training interviews ranged from 0 to 10 ($M = 3.37$, $SD = 3.62$).

Interview length

The average length of the interviews was 30 minutes and 21 seconds ($SD = 5$ minutes and 51.46 seconds). A 2 (Training) × 2 (Veracity) × 2 (Scenario) analysis of variance (ANOVA) with length of interview as a dependent variable revealed a scenario main effect, $F(1, 42) = 4.60$, $p = .038$, $d = 0.68$. All other effects were not significant (all F's < 1.62, all p's $> .21$). The secret meeting interviews ($M = 32$ minutes and 22 seconds, $SD = 4$ minutes and 3.91 seconds, 95% confidence interval [CI] [29 minutes and 41.49 seconds, 34 minutes and 27. 88 seconds]) were longer than the restaurant interviews ($M = 28$ minutes and 50 seconds, $SD = 6$ minutes and 48. 34 seconds, 95% CI [26 minutes and 6.36 seconds, 30 minutes and 52.75 seconds]).

RESULTS

Manipulation checks

Four 2 (Training) × 2 (Veracity) × 2 (Scenario) ANOVAs were conducted on the four manipulation checks. The analysis for motivation did not reveal any significant effects (all F's < 1.01, all p's $> .32$). The average motivation score was very high ($M = 6.16$, $SD = 0.83$ on a 7-point Likert scale). The analysis for the likelihood of receiving an incentive revealed a main effect for veracity, $F(1, 42) = 4.70$, $p = .036$, $d = 0.64$, with truth tellers judging the likelihood as higher ($M = 5.25$, $SD = 1.42$, 95% CI [4.65, 5.87]) than liars

DOI: 10.1002/jip

($M = 4.35$, $SD = 1.41$, 95% CI [3.77, 4.93]). All other effects were not significant (all F's < 0.79, all p's $> .38$). The analysis for the likelihood of receiving a penalty did not reveal a significant effect (all F's < 2.32, all p's $> .135$). Finally, the analysis about levels of cognitive load experienced resulted in a main effect for veracity, $F(1, 42) = 3.98$, $p = .053$, $d = 0.61$, with liars ($M = 5.12$, $SD = 1.53$, 95% CI [3.54, 4.89]) experiencing more cognitive load than truth tellers ($M = 4.26$, $SD = 1.27$, 95% CI [2.77, 4.19]). All other effects were not significant (all F's < 1.49, $p > .23$).

Police officers' accuracy

A 2 (Training) × 2 (Veracity) × 2 (Scenario) ANOVA with accuracy as detail revealed significant main effects for veracity, F(1, 42) = 15.55, $p < .001$, $d = 1.50$, and a significant Training × Veracity, $F(1, 42) = 4.47$, $p = .018$, $\eta^2 = .18$ interaction effect. All other effects were not significant, all F's < 1.21, all p's $> .30$, including the difference in accuracy obtained by untrained ($M = 0.59$, $SD = 0.50$, 95% CI [0.40, 0.78]) and trained detectives ($M = 0.65$, $SD = 0.49$, 95% CI [0.45, 0. 68]), $F(1, 42) = 0.95$, $p = .335$, $d = 0.12$. Truth accuracy ($M = 0.92$, $SD = 0.28$, 95% CI [0.75, 1.05]) was higher than lie accuracy ($M = 0.35$, $SD = 0.48$, 95% CI [0.20, 0.49]). The truth accuracy rate was significantly above chance ($t(23) = 7.23$, $p < .001$, $d = 1.50$) and significantly above the average truth accuracy obtained by professionals in 30 lie detection samples (56%, Vrij, 2008, $t(23) = 6.19$, $p < .001$, $d = 1.28$). The lie accuracy rate did not differ from chance ($t(23) = 1.62$, $p = .12$, $d = 0.31$) but was significantly below the average lie accuracy obtained by professionals in 30 lie detection samples (56%, Vrij, 2008, $t(25) = 2.25$, $p = .03$, $d = 0.44$).

The significant Training × Veracity interaction effect revealed that truth accuracy was similar for untrained ($M = 1.00$, $SD = 0$, 95% CI [0.85, 1.15]) and trained detectives ($M = 0.80$, $SD = 0.42$, 95% CI [0.62, 0.98]), $F(1, 22) = 3.21$, $p = .087$, $d = 0.95$, but that lie accuracy significantly increased from $M = 0.15$ ($SD = 0.37$, 95% CI [−0.11, 0.41]) for untrained detectives to $M = 0.54$ ($SD = 0.13$, 95% CI [0.28, 0.80]) for trained detectives, $F(1, 24) = 4.69$, $p = .041$, $d = 1.56$. This latter finding gives limited support to Hypothesis 1. The lie accuracy rate obtained by the untrained detectives was significantly below chance, $t(12) = 3.32$, $p = .006$, $d = 0.95$, and significantly below the 56% accuracy rate obtained on average in deception studies with professionals, $t(12) = 3.90$, $p = .002$, $d = 1.11$, whereas the lie accuracy rate obtained by the trained detectives was at chance level, $t(12) = 0.267$, $p = .794$, $d = 0.31$, and did not differ significantly from the 56% accuracy rate obtained on average in deception studies with professionals, $t(12) = 0.15$, $p = .88$, $d = 0.15$.

Detail elicited in the interviews

A 2 (Training) × 2 (Veracity) × 2 (Scenario) ANOVA with detail as a dependent variable resulted in a main effect for Training, $F(1, 42) = 4.55$, $p = .39$, $d = 0.60$, and Scenario, $F(1, 42) = 23.13$, $p < .001$, $d = 1.53$, with all other effects being not significant (all F's < 1.36, all p's $> .25$). Untrained detectives elicited fewer detail ($M = 147.66$, $SD = 57.19$, 95% CI [122.75, 176.02]) than trained detectives ($M = 195.59$, $SD = 103.05$, 95% CI [162.01, 220.17]). The restaurant scenario elicited more detail ($M = 217.08$, $SD = 94.44$, 95% CI [189.34, 245.11]) than the secret meeting scenario ($M = 122.34$, $SD = 29.59$, 95% CI [95.36, 151.13]).

The fact that we found no difference in detail between truth tellers and liars contradicts the general trend in deception research that truth tellers provide more details than liars

DOI: 10.1002/jip

(DePaulo *et al*., 2003; Vrij, 2005, 2008). This was also the case in a recent study in which the secret meeting scenario was used (Shaw *et al*., 2014). A possible reason is that the officers asked many questions that were irrelevant for lie detection purposes. For example, a key difference between truth tellers and liars was that truth tellers stayed for about 15 minutes in the restaurant, whereas liars did not. Detectives asked many questions about what happened before they entered the restaurant (33% of the questions were about this topic), but they are irrelevant for lie detection purposes because liars' and truth tellers' activities did not differ from each other before they entered the restaurant. We carried out further analyses and only included the detail elicited by questions about 'within the restaurant' (57% of questions). A 2 (Training) × 2 (Veracity) ANOVA with those detail as dependent variable revealed no significant effects either, all F's < 2.82, all p's > .10. The absence of a difference between truth tellers and liars means that Hypothesis 2 was rejected.

Police detectives' questions

A 2 (Training) × 2 (Veracity) × 2 (Scenario) ANOVA with the number of questions asked as a dependent variable revealed a main effect for training, $F(1, 42) = 9.77$, $p = .003$, $d = 0.93$, whereas all other effects were not significant, all F's < 2.75, all p's > .10. The untrained detectives asked more questions ($M = 98.71$, $SD = 36.00$, 95% CI [86.29, 111.12]) than the trained detectives ($M = 69.59$, $SD = 26.72$, 95% CI [56.14, 83.04]).

A 2 (Training) × 2 (Veracity) × 2 (Scenario) ANOVA with the appropriateness of the questions as a dependent variable revealed a training main effect, $F(1, 42) = 4.21$, $p = .046$, $d = 0.60$, whereas all other effects were not significant, all F's < 3.39, all p's > .07. A significantly higher proportion of appropriate questions was asked by trained detectives ($M = 0.42$, $SD = 0.12$, 95% CI [0.38, 0.46]) than by untrained detectives ($M = 0.36$, $SD = 0.08$, 95% CI [0.32, 0.40]), which supports Hypothesis 3.

Both the percentages obtained by untrained, $t(26) = 4.27$, $p < .001$, $d = 0.875$, and trained detectives, $t(26) = 5.30$, $p < .001$, $d = 1.08$, were significantly higher than the 29% obtained by Oxburgh *et al*. (2012). Of the three types of appropriate questions (open-ended questions, probing questions, and encourager/acknowledgements), only open-ended questions yielded a significant difference between untrained and trained detectives, $F(1, 42) = 7.94$, $p = .007$, $d = 0.81$ (for the other two categories, both F's < 1.07, p's > .30). Untrained detectives asked fewer open-ended questions ($M = 0.067$, $SD = 0.039$, 95% CI [0.047, 0.086]) than trained detectives ($M = 0.107$, $SD = 0.06$, 95% CI [0.086, 0.128]).

The trained police detectives indicated that they had used a variety of the techniques taught in the training in their interviews, with asking interviewees to draw being the most frequently used technique (Table 1). (We did not ask the untrained detectives this question.) We coded the transcripts for the presence of four of the taught techniques (model statement, reverse order, drawings, and spatial questions). A 2 (Training) × 2 (Veracity) × 2 (Scenario) ANOVA with the number of techniques taught as dependent variable revealed a main effect for training, $F(1, 42) = 26.51$, $p < .001$, $d = 1.86$. All other effects were not significant, all F's < 0.49, all p's > .48. A higher proportion of techniques taught in the training were used by trained detectives ($M = 0.15$, $SD = 0.15$, 95% CI [0.11, 0.19]) than by untrained detectives ($M = 0.002$, $SD = 0.009$, 95% CI [−0.04, 0.04]). Finally, we correlated the proportion of questions based on the taught-with (total) accuracy in discriminating between truth tellers and liars (after-training interviews only). The correlation was not significant, $r(23) = 0.01$, $p = 0.96$, which means that Hypothesis 4b is rejected.

DOI: 10.1002/jip

In fact, one untrained detective asked one of the question introduced in the training, whereas 20 trained detectives did at least ask one of the questions introduced in the training in their interviews. In contrast, all 23 trained detectives thought themselves to have used at least on of these four taught techniques (model statement, reverse order, drawings, and spatial questions) in their interviews. Our analyses revealed that, of the four techniques we coded for in the transcripts, the detectives used on average $M = 1.78$ ($SD = 1.04$, 95% CI [1.48, 2.12]) techniques, whereas the trained detectives themselves thought they used significantly more of these techniques ($M = 2.78$, $SD = 1.12$, 95% CI [2.30, 3.27]), $t(23) = 5.30$, $p < .001$, $d = 93$.

The 20 trained detectives mentioned earlier who correctly incorporated the taught techniques in their interviews elicited significantly more detail with questions based on the techniques taught in the training ($M = 5.79$ details per question, $SD = 3.90$, 95% CI [3.90, 7.67]) than with the other questions they asked ($M = 2.31$, $SD = 2.02$, 95% CI [1.33, 3.28]), $t(19) = 3.43$, $p = .003$, $d = 1.18$, which supports Hypothesis 4a.

DISCUSSION

Summary of the findings

The training increased the percentage of appropriate questions asked by the police detectives in their interviews. This increase was entirely due to an increase in open-ended questions, which was expected as open-ended questions were used in the training. The percentage of appropriate questions after training (42%) was significantly above the percentage (29%) obtained by Oxburgh *et al*. (2012) in their analysis of interviews by English (and Welsh) police detectives, and 42% is a substantial improvement compared with 29%. This means that our sample did relatively well.

We further found that the questions related to four of the taught techniques (model statement, reverse order, drawing, and spatial questions) elicited more detail ($M = 5.79$ on average) than the other questions the police detectives asked ($M = 2.31$ on average). This is an important finding as eliciting detail is at the core of investigative interviewing (Bull, 2010; Fisher, 2010). Apparently, the techniques taught in the training contributed to this core aspect. Not many questions asked in the interviews were related to the taught techniques (15%), but this percentage looks worse than it actually is. Many questions related to the taught techniques are open-ended questions that invite free recall of the entire event (e.g. 'Please tell me in as much detail as possible what you saw in the video and start by describing what you saw at the end of the video'), and there are only a limited number of such questions that can be asked in a single interview.

The training had little effect on the detectives' ability to distinguish truth tellers from liars. No correlation was found between asking questions taught in the training and ability to discriminate between truth tellers and liars. Accuracy in lie detection significantly increased from 15% to 54% as a result of training, whereas no difference was found in truth accuracy between before and after training. However, the significant increase in lie detection is not really meaningful. First, although lie accuracy improved as a result of training, the accuracy level was not above a level that could be expected by chance alone. Second, although truth accuracy rate did not differ between without and after training interviews, the percentage went down from 100% without training to 80% after training. The truth and lie accuracy percentages combined show that untrained detectives showed a strong truth bias (nearly all of the judgments they made were truth judgements) and that this truth

DOI: 10.1002/jip

bias diminished somewhat after training. A weakened truth bias (or, rather, increased lie bias) is a known effect of training when the training concentrates on 'cues to deceit' (Masip, Alonso, Garrido, & Herrero, 2009). The opposite, an increase in truth bias, occurs when the training focuses on 'cues to truthfulness' (Masip *et al*., 2009). The present training focussed on cues to deceit and on cues to truthfulness and should therefore, in theory, not show any change in bias. However, a weakened truth bias was always likely in this experiment given the strong truth bias and 100% truth accuracy obtained by the untrained police detectives. This ceiling effect means that the trained police detectives could never improve upon this truth accuracy rate. As such, the experiment is not a proper test of the training workshop's potential to increase truth accuracy, and a setting in which the untrained group obtains more modest accuracy rates would allow more room to demonstrate the true potential of the training.

Regarding the amount of detail elicited from interviewees, this is one of only two deception experiments in which police detectives interviewed mock suspects and in which it was examined how many details they elicited from truth tellers and liars. Truth tellers and liars provided a similar amount of detail. This replicates the findings of Vrij *et al*. (2015)—the second study in which this issue was investigated—but it goes against the general trend in deception research where it is typically found that truth tellers provide more detail than liars (DePaulo *et al*., 2003; Masip *et al*., 2005; Vrij, 2005, 2008). This was also the case in a recent study in which the secret meeting scenario was used (Shaw *et al*., 2014).

We believe that the questions asked by the police detectives are responsible for the absence of difference in detail between truth tellers and liars and the absence of a training effect on accuracy rates. In the present experiment, reflecting real life (Leins, Fisher, & Ross, 2013), mock suspects did not have to lie to each question and, for example in the restaurant scenario, could have answered questions truthfully about the person who brought them to the restaurant and the layout of the restaurant. Analyses of the content of the questions revealed that the police detectives asked many questions about the confederate and about his interactions with the mock suspects, and these questions are unlikely to reveal information about deceit. Even when we focussed on the part of the scenario in which truth tellers and liars did differ (what happened inside the restaurant), differences in detail between truth tellers and liars did not emerge, perhaps again due to asking the wrong questions. For example, some detectives asked participants to describe in detail the layout of the restaurant, but because liars also had been in the restaurant, they would also be able to do this, negating possible differences between truth tellers and liars.

A possible explanation for asking—for lie detection purposes—the wrong questions is the lack of training in the UK in lie detection. Reading police interview manuals suggests that in the USA lie detection plays an important role in training (Inbau, Reid, Buckley, & Jayne, 2013)—albeit of poor quality according to experimental research (e.g. Kassin & Fong, 1999; Mann, Vrij, & Bull, 2004)—but, as our trainees told us, lie detection training does not take place in the UK. The trainees reported that they have been taught that cues to deceit are generally unreliable (a statement backed up by research) and that they are instructed to focus on eliciting information instead. We believe that this is a shortcoming in UK police training as a substantial body of research has shown that specific questioning techniques, including the ones taught in the training, result in the elicitation of detail and facilitate truth/lie detection.

A second reason why the detectives often did not ask the right questions is that a 1-day training, in all likelihood, is too short to yield strong positive effects (Hauch *et al*., 2014). A 1-day training course in interviewing with limited time to practice and without face-to-face

DOI: 10.1002/jip

sessions is not an ideal way to teach people new skills (Powell *et al*., 2005), and our training may not have long-lasting effects. However, for long-lasting effects to develop, short-term effects should be evident first, and the present training study showed positive effects in terms of asking appropriate questions. The police training literature warns about the difficulties of teaching investigative interview techniques and to make sure that trainees subsequently introduce them into their work practice (see Powell *et al*., 2005, for a review of that research). We clearly experienced such difficulties. Although the police detectives thought that they had implemented our techniques in the way we taught them to do, this often did not appear to be the case. This means that considerable attention needs to be paid on how to implement taught techniques into practice for which in-group and face-to-face sessions need to be held that incorporate demonstrations, exercises, and feedback (Powell *et al*., 2005). Another challenge is to ensure that trainees will not fall back to their common practice over time, which can only be overcome by having refreshment training courses periodically (Powell *et al*., 2005). Thus, a 1-day training in cognitive lie detection is an important first step in making investigators aware of the complexity, difficulties, and opportunities in lie detection, but a training lasting multiple days followed up by periodic refreshment trainings is probably required to achieve its full potential.

A comparison between two training studies

Vrij *et al*. (2015) reported a training study similar to the training study reported in the present article. The only difference was that the training in Vrij *et al*. (2015) was delivered by the scientists who developed the training whereas the interview techniques used in the training discussed in the present article were introduced by a retired practitioner. The two studies produced several similar findings. Training significantly increased the percentage of appropriate questions asked in the interview, and the percentage of appropriate questions achieved in Vrij *et al*. (2015) was similar (40%) to the percentage obtained in the present study (42%). The percentage of taught questions asked in the interviews was a bit lower (7%) in Vrij *et al*. (2015) compared with the present study (15%), but the number of details elicited per questions was a bit higher ($M = 10.92$) in Vrij *et al*. (2015) than in the present study ($M = 5.79$). Both studies revealed similar challenges in that the detectives had difficulty in implementing the techniques in the way it was taught to them and that their questioning elicited the same number of detail in truth tellers and liars. As we discussed earlier, the fact that the training lasted only 1 day and the subsequent lack of opportunity to practice the techniques may have caused these null effects. The only noticeable difference between the two studies was that in Vrij *et al*. (2015) the training increased the ability to distinguish between truth tellers and liars whereas this was not the case in the present experiment. Our data cannot explain this difference in results, but we think that the ability to detect truths and lies will increase if the 1-day training workshop will be replaced by a multiple-day training workshop.

Methodological considerations

We finish this article with some methodological considerations. In this study, truth tellers and liars were given detailed instructions. In the restaurant scenario, truth tellers were instructed to report that they spent time in the restaurant (rather than being given the vaguer instruction 'to demonstrate their innocence'). We gave this explicit instruction because we were afraid that not all truth tellers would understand what to do if they were simply asked to 'demonstrate

DOI: 10.1002/jip

their innocence'. Participants who fail to understand the instructions produce experimental noise. Liars were instructed to say that they had been in the restaurant. This instruction was given so that we could compare the truth tellers' and liars' stories and that it would not be obvious to the detective who would be lying (without such an instruction, liars who denied having been to the restaurant would then be easily and correctly classified as liars).

In the secret meeting scenario, participants were instructed to be 'completely' truthful. In intelligence settings, it is often desirable that agents from the same agency are entirely truthful to each other, and we wanted to reflect that situation. Of course, this does not imply that all truth tellers told the interviewers all they remembered. Research has demonstrated that cooperative witnesses often do not provide all the information they remember (Vrij *et al*., 2014). Liars were instructed to tell a mixture of truths and lies, which reflects real life in which it rarely happens that liars tell a completely fabricated story (Leins *et al*., 2013). The rationale given to the participants to provide truthful information was that it was important to appear cooperative and that the investigator had some information about the meeting they had witnessed. Such a situation (the interviewee's task is to appear cooperative, and the interviewer may have information about the topic of investigation) is common in intelligence interviews (Vrij & Granhag, 2014). Of course, this means that we gave liars in the secret meeting scenario a difficult task to complete. However, the difficulty of the truth tellers' task should not be underestimated either, as they were instructed to remember and subsequently report back details of a meeting in which a large amount of information was provided.

The truth accuracy rates (100% without training and 80% with training) were very high and much higher than typically obtained in deception research. We do not think that this is caused by the experimental scenarios used in the training, as they are typical for deception research and used by us before. We believe that the sample (UK police detectives) is responsible for the effect. Some officers told us that they prefer to refrain from accusing someone of being a liar without being certain about it. We believe that this is an admirable approach.

We instructed trained police to implement, where possible, some of the techniques they had learnt during the training in their interviews. We did not give them a stricter instruction, such as 'implement as many techniques as you can in the interview'. We were afraid that this would make the interviewing artificial, something we wanted to avoid. Of course, this means that we do not know why the detectives did not use all the techniques: Did they did not comprehend them sufficiently to implement them, or did they think it was not appropriate or necessary to implement them?

Finally, although the practitioner had a vital role in the training (by introducing and explaining the techniques), scientists were still present in the training to support the practitioner. As such, this study was only the first step in a science-based training delivered by practitioners rather than the final product. Future work should examine whether a science-based training can be delivered entirely by practitioners. This is an important—but often neglected—aspect of applied science, because only when practitioners can deliver the training is there a chance that such a training can be delivered on a wide scale.

ACKNOWLEDGEMENT

This work is funded by the High-Value Detainee Interrogation Group/J-FBI-10-009 awarded to the University of Portsmouth (UK). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the U.S. Government.

DOI: 10.1002/jip

REFERENCES

Bell, B. E., & Loftus, E. F. (1989). Trivial persuasion in the courtroom: The power of (a few) minor details. *Journal of Personality and Social Psychology*, *56*, 669–679. DOI:10.1037//0022-3514.56.5.669.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgements. *Personality and Social Psychology Review*, *10*, 214–234. DOI:10.1207/s15327957pspr1003_2.

Bull, R. (2010). The investigative interviewing of children and other vulnerable witnesses: Psychological research and working/professional practice. *Legal and Criminological Psychology*, *15*, 5–24. DOI:10.1348/014466509X440160.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex*, *19*, 1557–1566. DOI:10.1093/cercor/bhn189.

Debey, E., Verschuere, B., & Crombez, G. (2012). Lying and executive control: An experimental investigation using ego depletion and goal neglect. *Acta Psychologica*, *140*, 133–141. DOI:10.1016/j.actpsy.2012.03.004.

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

Evans, J. R., Meissner, C. A., Michael, S. W., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a psychologically based credibility assessment tool. *Journal of Applied Research in Memory and Cognition*, *2*, 33–41. DOI:10.1016/j.jarmac.2013.02.002.

Farah, M. J., Hutchinson, J. B., Phelps, E. A., & Wagner, A. D. (2014). Functional MRI-based lie detection: Scientific and societal challenges. *Nature Reviews Neuroscience*, *15*, 123–131. DOI:10.1038/nrn3665.

Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology*, *15*, 25–38. DOI:10.1348/135532509X441891.

Fisher, R. P., Falkner, K. L., Trevisan, M., & McCauley, M. R. (2000). Adapting the cognitive interview to enhance long-term (35 years) recall of physical activities. *Journal of Applied Psychology*, *85*, 180–189. DOI:10.1037/0021-9010.85.2.180.

Granhag, P. A., & Hartwig, M. (2015). The strategic use of evidence (SUE) technique: A conceptual overview. In Granhag, P. A., Vrij, A., & Verchuere, B. (Eds.), Deception detection: Current challenges and new approaches. Chichester, England: Wiley.

Griffiths, A., & Milne, R. (2006). Will it al end in tiers? Police interviews with suspects in Britain. In T. Williamson (Ed.), Investigative interviewing: Rights, research, regulation (pp. 167–189). Cullompton: Willan Publishing.

Gudjonsson, G. H., & Clark, N. K. (1986). Suggestibility in police interrogation: A social psychological model. *Social Behaviour*, *1*, 83–104.

Hartwig, M., Granhag, P. A., & Luke, T. (2014). Strategic use of evidence during investigative interviews: The state of the science. In: D. C. Raskin, C. R. Honts, & J. C. Kircher (Eds.), Credibility assessment: Scientific research and applications (pp. 1–36). Oxford, UK: Academic Press.

Hartwig, M., Granhag, P. A., & Strömwall, L. (2007). Guilty and innocent suspects' strategies during interrogations. *Psychology, Crime & Law*, *13*, 213–227. DOI:10.1080/10683160600750264.

Hartwig, M., Granhag, P. A., Strömwall, L., & Kronkvist, O. (2006). Strategic use of evidence during police interrogations: When training to detect deception works. *Law and Human Behavior*, *30*, 603–619. DOI:10.1007/s10979-006-9053-9.

Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2014). Does training improve the detection of deception? A meta-analysis. *Communication Research*, published online 25 May 2014, doi: 10.1177/0093650214534974

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2013). Criminal interrogation and confessions (5th ed. pp.). Burlington, MA: Jones & Bartlett Learning.

Johnson, M. K. (2006). Memory and reality. *American Psychologist*, *61*, 760–771. DOI:10.1037/0003-066X.61.8.760.

Johnston, W. A., Greenberg, S. N., Fisher, R. P., & Martin, D. W. (1970). Divided attention: A vehicle for monitoring memory processes. *Journal of Experimental Psychology*, *83*, 164–171. DOI:10.1037/h0028554.

DOI: 10.1002/jip

Kassin, S. M., & Fong, C. T. (1999). 'I'm innocent!': Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior*, *23*, 499–516. DOI:10.1023/A:1022330011811.

Leal, S., Vrij, A., Warmelink, L., Vernham, Z., & Fisher, R. (2015). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology*, *20*, 129–146. DOI:10.1111/lcrp.12017.

Leins, D., Fisher, R. P., & Ross, S. J. (2013). Exploring liars' strategies for creating deceptive reports. *Legal and Criminological Psychology*, *18*, 141–151. DOI:10.1111/j.2044-8333.2011.02041.x.

Leins, D., Fisher, R., & Vrij, A. (2012). Drawing on liars' lack of cognitive flexibility: Detecting deception through varying report modes. *Applied Cognitive Psychology*, *26*, 601–607. DOI:10.1002/acp.2837.

Leins, D., Fisher, R. P., Vrij, A., Leal, S., & Mann, S. (2011). Using sketch-drawing to induce inconsistency in liars. *Legal and Criminological Psychology*, *16*, 253–265. DOI:10.1348/135532510X501775.

Lisofsky, N., Kazzer, P., Heekeren, H. R., & Prehn, K. (2014). Investigating socio-cognitive processes in deception: A quantitative meta-analysis of neuroimaging studies. *Neuropsychologia*, *61*, 113–122. DOI:10.1016/j.neuropsychologia.2014.06.001.

Mac Giolla, E., Granhag, P. A., & Liu-Jönsson, M. (2013). Markers of good planning behaviour as a cue for separating true and false intent. *PsyCh Journal*, *2*, 183–189. DOI:10.1002/pchj.36.

Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect deceit. *Journal of Applied Psychology*, *89*, 137–149. DOI:10.1037/0021-9010.89.1.137.

Mann, S., Vrij, A., Shaw, D., Leal, S., Ewens, S., Hillman, J., … Fisher, R. P. (2013). Two heads are better than one? How to effectively use two interviewers to elicit cues to deception. *Legal and Criminological Psychology*, *18*, 324–340. doi: 10.1111/j.2044-8333.2012.02055.x

Masip, J., Alonso, H., Garrido, E., & Herrero, C. (2009). Training to detect what? The biasing effects of training on veracity judgements. *Applied Cognitive Psychology*, *23*, 1282–1296. DOI:10.1002/acp.1535.

Masip, J., Sporer, S., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law*, *11*, 99–122. DOI:10.1080/1068316041000172635.

Myklebust, T., & Bjorklund, R. A. (2006). The effect of long-term training on police officers' use of open and closed questions in field investigative interviews of children (FIIC). *International Journal of Investigative Psychology and Offender Profiling*, *3*, 165–181. DOI:10.1002/jip.52.

Nahari, G., Vrij, A., & Fisher, R. P. (2014). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology*, *19*, 227–239. DOI:10.1111/j.2044-8333.2012.02069.x.

Oxburgh, G., & Dando, C. J. (2011). Psychology and interviewing: What direction now in our quest for reliable information. *The British Journal of Forensic Practice*, *13*, 135–144. DOI:10.1108/14636641111134378.

Oxburgh, G., Ost, J., & Cherryman, J. (2012). Police interviews with suspected child sex offenders: Does use of empathy and question type influence the amount of investigation relevant information obtained? *Psychology, Crime & Law*, *18*, 259–273. DOI:10.1080/1068316X.2010.481624.

Powell, M. B., Fisher, R. P. & Wright, R. (2005). Investigative interviewing. In N. Brewer, & K. D. Williams (Eds.), Psychology and law: An empirical perspective (pp. 11–42). New York: Guilford Press.

Roos af Hjelmsäter, E., Öhman, L., Granhag, P. A., & Vrij, A. (2014). Mapping deception in adolescents: Eliciting cues to deceit through an unanticipated spatial drawing task. *Legal and Criminological Psychology*, *19*, 179–188. DOI:10.1111/j.2044-8333.2012.02068.x.

Shaw, D., Vrij, A., Leal, S., & Mann, S. (2014). 'We'll take it from here': The effect of changing interviewers in information gathering interviews. *Applied Cognitive Psychology*, *28*, 908–916. DOI:10.1002/acp.3072.

Shaw, D. J., Vrij, A., Leal, S., Mann, S., Hillman, J., Granhag, P. A., & Fisher, R. P. (2015, published online). Mimicry and investigative interviewing: Using deliberate mimicry to elicit information and cues to deceit. *Journal of Investigative Psychology and Offender Profiling*. doi: 10.1002/jip.1438

Shaw, D. J., Vrij, A., Mann, S., Leal, S., & Hillman, J. (2013). Expect the unexpected? Variations in question type elicit cues to deception in joint interviewer contexts. *Applied Cognitive Psychology*, *27*, 336–343. DOI:10.1002/acp.2911.

Snook, B., Luther, K., Quinlan, H., & Milne, R. (2012). Let 'em talk! A field study of police questioning practices of suspects and accused persons. *Criminal Justice and Behavior*, *39*, 1328–1339. DOI:10.1177/0093854812449216.

Sporer, S. L. (2004). Reality monitoring and detection of deception. In P. A. Granhag, & L. A. Strömwall (Eds.), Deception detection in forensic contexts (pp. 64-102). Cambridge, England: Cambridge University Press.

Steller, M., & Köhnken, G. (1989). Criteria-based content analysis. In D. C. Raskin (Ed.), Psychological methods in criminal investigation and evidence (pp. 217–245). New York: Springer-Verlag.

Strömwall, L. A., Granhag, P. A., & Jonsson, A. C. (2003). Deception among pairs: 'Let's say we had lunch together and hope they will swallow it'. *Psychology, Crime & Law*, *9*, 109–124. DOI:10.1080/1068316031000116238.

Vernham, Z., Vrij, A., Mann, S., Leal, S., & Hillman, J. (2014). Collective interviewing: Eliciting cues to deceit using a turn-taking approach. *Psychology, Public Policy, and Law*, *20*, 309–324. DOI:10.1037/law0000015.

Visu-Petra, G., Varga, M., Miclea, M., & Visu-Petra, L. (2013). When interference helps: Increasing executive load to facilitate deception detection in the concealed information test. *Frontiers in Psychology*, *4*, 146. DOI:10.3389/fpsyg.2013.00146.

Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills and physical characteristics on impression formation. *Journal of Social Psychology*, *133*, 601–611.

Vrij, A. (1994). The impact of information and setting on detection of deception by police detectives. *Journal of Nonverbal Behavior*, *18*, 117–137. DOI:10.1007/BF02170074.

Vrij, A. (1995). Behavioral correlates of deception in a simulated police interview. *Journal of Psychology: Interdisciplinary and Applied*, *129*, 15–29. DOI:10.1080/00223980.1995.9914944.

Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, *11*, 3–41. DOI:10.1037/1076-8971.11.1.3.

Vrij, A. (2008). Detecting lies and deceit: Pitfalls and opportunities(2nd ed. pp.). Chichester: John Wiley and Sons.

Vrij, A. (2015). A cognitive approach to lie detection. In Granhag, P. A., Vrij, A., & Verchuere, B. (Eds.), Deception detection: Current challenges and new approaches. Chichester, England: Wiley.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences*, *10*, 141–142. DOI:10.1016/j.tics.2006.02.003.

Vrij, A., & Ganis, G. (2014). Theories in deception and lie detection. In D. C. Raskin, C. R. Honts, J. C. Kircher (Eds.), Credibility assessment: Scientific research and applications (pp. 301–374). Oxford, UK: Academic Press.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, *1*, 110–117. DOI:10.1016/j.jarmac.2012.02.004.

Vrij, A., & Granhag, P. A. (2014). Eliciting information and detecting lies in intelligence interviewing: An overview of recent research. *Applied Cognitive Psychology*, *28*, 936–944. DOI:10.1002/acp.3071.

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Towards a cognitive lie detection approach. *Current Directions in Psychological Science*, *20*, 28–32. DOI:10.1177/0963721410391245.

Vrij, A., Granhag, P. A., & Porter, S. B. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, *11*, 89–121. DOI:10.1177/1529100610390861.

Vrij, A., Hope, L., & Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights from Behavioral and Brain Sciences*, *1*, 129–136. DOI:10.1177/2372732214548592.

Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J., & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior*, *33*, 159–166. DOI:10.1007/s10979-008-9143-y.

Vrij, A., Leal, S., Mann, S., Vernham, Z., & Brankaert, F. (2015). Translating theory into practice: Evaluating a cognitive lie detection training workshop. Manuscript under review.

Vrij, A., Leal, S., Mann, S., Warmelink, L., Granhag, P. A., & Fisher, R. P. (2010). Drawings as an innovative and successful lie detection tool. *Applied Cognitive Psychology*, *4*, 587–594. DOI:10.1002/acp/1627.

DOI: 10.1002/jip

Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, *32*, 253–265. DOI:10.1007/s10979-007-9103-y.

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2012). Is anyone there? Drawings as a tool to detect deception in occupations interviews. *Psychology, Crime & Law*, *18*, 377–388. DOI:10.1080/1068316X.2010.498422.

Walsh, D., & Bull, R. (2012). Examining rapport in investigative interviews with suspects: Does its building and maintenance work? *Journal of Police and Criminal Psychology*, *27*, 73–84. DOI:10.1007/s11896-011-9087-x.

APPENDIX 1
Information about the Training Workshop and Procedure

Demonstrations, videos, and exercises used in the training workshop

Demonstrations

- A demonstration showing that hard thinking leads to a decrease in movements (Pitfalls section).
- A demonstration showing that people are typically unaware of being mimicked (Opportunities section).

Videos

- Videos showing that Bill Clinton blinked less when he lied compared with telling the truth (Pitfalls section).
- Videos showing that in specific interview situations liars tend to move less than truth tellers (Pitfalls section). Videos based on Vrij's (1993, 1994, 1995) research paradigm.
- Videos showing the effect of turn-taking on truth-telling and lying pairs (Opportunities section). Videos based on the research paradigm of Vernham *et al*. (2014).
- Videos showing the effect of using a model statement on the amount of information given by truth tellers and liars (Opportunities section). Videos based on the research paradigm of Leal *et al*. (2015).

Exercises

- A lie detection test by paying attention to non-verbal behaviour. Two trainees took part as liar and truth teller; the others were observers (Pitfalls section).
- A lie detection test by using the model statement. Two trainees took part as liar and truth teller; the others were observers (Opportunities section).
- A lie detection test based on observing drawings. The drawings were taken from Vrij *et al*. (2010), and the trainees were observers (Opportunities section).

DOI: 10.1002/jip