An Empirical Test of the Behaviour Analysis Interview

Aldert Vrij,[1,3] **Samantha Mann,**[1] **and Ronald P. Fisher**[2]

The present experiment is the first empirical test of the Behaviour Analysis Interview (BAI), an interview technique developed by F. E. Inbau, J. E. Reid, J. P. Buckley, & B. C. Jayne (2001) designed to evoke different verbal and non-verbal responses from liars and truth-tellers. Inbau et al. expect liars to be less helpful than truth-tellers in investigations and to exhibit more nervous behaviours. Just the opposite predictions, however, follow from the deception literature, which notes that liars take their credibility less for granted and are therefore more aware of their responses and their impact on others. This suggests that liars' answers should be more helpful than truth-tellers' answers, and liars' non-verbal responses should appear more relaxed than truth-tellers' non-verbal responses. In the present experiment, 40 participants (undergraduate students) lied or told the truth about an event during a BAI interview. The interviews were coded according to Inbau et al.'s guidelines. The results showed that, compared to liars, truth-tellers (a) were more naive and evasive when explaining the purpose of the interview, and (b) were less likely to name someone who they felt certain did not commit the crime. Truth-tellers also exhibited more nervous behaviours. The results were consistent with the predictions of the deception literature, and directly opposed to the predictions of BAI.

Scientists, law enforcement agents and a host of other investigators have been attempting to develop better methods to discriminate between truth-tellers and liars. Towards this goal, researchers have examined respondents ' physiological responses (e.g., the control question polygraph test and the guilty knowledge polygraph test, see Kleiner, 2002) and verbal responses (e.g., Statement Validity Analysis; Köhnken & Steller, 1988; Steller & Köhnken, 1989; Vrij, 2000). The greater the difference is between the physiological and verbal responses of truth-tellers and liars, the easier it is to discriminate between them. One might expect, therefore, that a comparable parallel line of research would have been developed to maximise the differences

[1] Psychology Department, University of Portsmouth, Portsmouth, UK.
[2] Florida International University, Miami Florida.
[3]To whom correspondence should be addressed at Psychology Department, University of Portsmouth, King Henry Building, King Henry 1 Street, Portsmouth PO1 2DY, UK; e-mail: aldert.vrij@port.ac.in.uk.

between truth-tellers' and liars' non-verbal responses. Perhaps there is a set of instructions or a unique task that we can give to respondents that will facilitate discriminating between the behaviour of truth-tellers and liars. This parallel line of research, unfortunately, has not emerged. For example, Ekman and his colleagues have demonstrated that paying attention to certain facial micro-expressions can be used to detect deceit (Ekman, 1985; Frank & Ekman, 1997); however, they provide no guidelines about how to conduct the interview in order to evoke these discriminating facial expressions. Other training programmes to detect deceit via non-verbal behaviour follow a similar pattern (Frank & Feeley, 2003; Vrij, 2000). Observers are informed about which behavioural cues they should focus on and which cues they should ignore to detect deceit, but they are given no guidance about which questions to ask to evoke these cues.

The only published interview protocol that we are aware of that is designed to evoke behavioural responses to detect lying is Inbau, Reid, Buckley, and Jayne's (2001) Behaviour Analysis Interview (BAI) technique. According to the BAI, truthful interviewees will respond differently to the specialised questions than will deceptive interviewees. We tested this claim in the present experiment. The BAI technique consists of 15 questions that are tailored towards the critical event in the current experiment, the theft of money from a wallet. An overview of these questions is given in the appendix. The bold labels are the labels that Inbau et al. (2001) gave to these questions, and we use their labels throughout this article. Blair (1998) introduced an additional 16th question, which is also listed in the appendix. In itself, the term BAI is misleading as BAI judges look at both non-verbal and verbal responses, and the questions are intended to evoke differences between truth-tellers and liars in both their non-verbal and verbal responses. In fact, when they describe their BAI technique, Inbau et al. (2001) concentrate more on the verbal differences between liars and truth-tellers than on the non-verbal differences. They predict that all 15 questions will elicit different verbal responses between liars and truth-tellers, whereas they examine non-verbal behaviours only in response to 3 of the 15 questions (Questions 2, 3 and 10).

Regarding the non-verbal responses, Inbau et al. (2001) assume that liars feel less comfortable than truth-tellers in the police interview situation. As a result, guilty suspects are more likely to cross their legs, shift in their chair, and perform grooming behaviours while answering the guilt question (Question 2), whereas innocent suspects are more likely to lean forward, establish eye contact and use illustrators to reinforce the confidence of their statements. While responding to the knowledge question (Question 3), guilty suspects are more likely to answer quickly. Their answers also will sound less sincere. Finally, during the response to the motive question (Question 10), guilty suspects are more likely to exhibit anxiety-reducing behaviours such as shifting posture in their chair.

Verbal responses should also discriminate between truth-tellers and liars, because, according to Inbau et al. (2001), interviewees' answers will reflect their different attitudes towards the investigation. Liars are thought to be less helpful in the interview and do not show the appropriate level of concern about being a suspect, whereas truth-tellers are more likely to offer helpful information, and show an expectancy to be exonerated (see also Horvath, Jayne, & Buckley, 1994). Thus,

compared to their innocent counterparts, guilty suspects are thought to be more evasive about the purpose of the interview (Question 1); less emphatic in their denial of having committed the crime (Question 2); more likely to deny any knowledge of who the culprit might be (Question 3); less likely to name another suspect (Question 4, because naming someone who the suspect knows is innocent would be an unnecessary lie); or to name someone they believe is innocent (Question 5, because guilty suspects prefer to surround themselves with other possible suspects). Guilty suspects are also less likely to admit that a crime has taken place (Question 6), and that they had the opportunity to commit the crime (Question 7). In addition, guilty suspects are more likely to voice negative feelings towards the interview (Question 8, because innocent suspects have faith that they will be exonerated), and are more likely to admit to having thought about committing a crime similar to that under investigation (Question 9, because guilty suspects will have an internal need to talk about their crimes in order to relieve anxiety, while at the same time escaping the consequences). Guilty suspects are also thought to be less likely to give a reasonable motive for the crime (Question 10, because guilty suspects do not want to reveal their own motives); less likely to suggest a serious punishment for the person who committed the crime (Question 11); and more likely to give such a person a second chance (Question 12). When asked why they would not commit the crime (Question 13), guilty suspects are believed to be more likely to answer in the third person ("That's against the law", "It's wrong"), whereas innocent suspects are more likely to answer in the first person ("Because I am not a thief"). Guilty suspects are also thought to express less confidence in being exonerated (Questions 14 and 16) and are less likely to have informed their loved ones about the interview (Question 15).

Inbau et al. (2001) state that research has demonstrated that innocent and guilty interviewees do respond differently to the BAI set of questions, but they do not give further details about this research. The only published work on BAI that we are aware of is Horvath et al.'s (1994) study in which police interviewed real criminal suspects. In this study, use of the technique resulted in 78% of the actual truthful suspects and 66% of the actual deceptive suspects being classified correctly. However, a problem with this study was establishing the ground truth, that is, determining with certainty the actual guilt or innocence of the suspects. Horvath et al. (1994) used as criteria to establish ground truth (i) confessions, and when these were not available, (ii) 'a systematic factual analysis' in which evaluators looked at factors such as 'biographical information', 'opportunity/access', and 'motivation'. Both of these criteria are problematic for establishing ground truth. Confessions are problematic because of the occurrence of false confessions (Gudjonsson, 2003; Kassin, 1997). Using a systematic factual analysis is perhaps even more problematic, because facts such as 'biographical information', 'opportunity' or 'motivation' do no provide reliable evidence about the actual guilt or innocence of a suspect. Horvath et al. (1994) themselves acknowledge the problems with using confessions or a systematic factual analysis as criteria to establish the ground truth. They state "In field settings it is extremely difficult to develop an adequate measure of ground truth" (p. 805). In only 2 of the 87 cases that they analysed was the ground truth established by 'incontrovertible evidence'. They concluded that, "If it were possible to develop ground truth criteria in a large number of cases such as occurred in these two instances,

the interpretation of findings would be less problematic" (p. 805). The result of a problematic ground truth is that it is unknown whether the alleged liars in this study were actually lying and the alleged truth-tellers were actually telling the truth. The present laboratory experiment is the first test of the efficiency of BAI where ground truth is certain.

Despite the apparent success of Horvath et al. (1994), there is good reason to doubt that Inbau's BAI hypothesis will promote good discrimination between truth-tellers and liars. The basic premise, that liars feel less comfortable than truth-tellers in a police interview, is not universally accepted by the scientific community. For instance, DePaulo et al. (2003) have argued that in situations where the consequences of not being believed are severe, not only liars but also truth-tellers will be concerned about not being believed. Similarly, Ekman (1985) has argued that both liars and truth-tellers feel uncomfortable when they are accused of serious wrongdoing.

Compatible with these last two views, but refuting Inbau's hypothesis, DePaulo et al.'s (2003) meta-analysis of non-verbal correlates of deception found that liars were not more likely than truth-tellers to look away, shift in their chair, cross their legs or make grooming gestures. In fact, eye contact is not related to deception, and, in direct contrast to Inbau et al.'s (2001) assumptions, liars tend to *de*crease rather than *in*crease their movements (DePaulo et al., 2003). The tendency of liars to decrease their movements may reflect greater cognitive load borne by liars. Lying is often cognitively more difficult then telling the truth (DePaulo et al., 2003; Ekman, 1985), and, as Ekman pointed out, greater cognitive load results in a neglect of body language, reducing overall animation (Ekman, 1997; Ekman & Friesen, 1972). Liars' decreased movements may be compatible with another account, namely impression management. Liars are likely to be concerned about the impression they make on others, even more so than truth-tellers (DePaulo & Kirkendol, 1989; DePaulo et al., 2003; Vrij, 2000; Vrij, Semin, & Bull, 1996). Like liars, truth-tellers are also keen to be seen as truthful. However, truth-tellers typically do not think that this will require any special effort or attention (DePaulo et al., 2003). In other words, truth-tellers more often take their credibility for granted (DePaulo, LeMay, & Epstein, 1991; DePaulo et al., 2003; Kassin & Norwick, 2004). Moreover, lying increases self-awareness. A liar has to act deliberately in order to avoid being caught, and in doing so, will be aware of his/her own behaviour and the impact it has on observers. Such awareness is not necessarily present in truth-tellers. As a result, liars are more inclined than truth-tellers to refrain from exhibiting excessive movements that could be construed as nervous or suspicious (DePaulo & Kirkendol, 1989).

Given the disparity between the lie detection principles espoused by Inbau et al. (2001) and the behavioural characteristics actually exhibited by liars and truth-tellers, one might even expect that adhering to the Inbau et al. principles may impair lie detection. In fact, Kassin and Fong (1999) found just that: People who were trained to look for the non-verbal cues to deception outlined by Inbau et al. performed significantly worse in a lie detection test than those who received no training (see also Meissner & Kassin, 2002). This finding is also compatible with a disturbing set of results reported by Mann, Vrij, and Bull (2004). In their lie

Table 1. Predictions According to the Behaviour Analysis Interview (BAI) and the Impression Management Hypothesis (IMH)

Question	Label	BAI Predictions	IMH predictions
Verbal cues			
Q1	Purpose	Liars are more evasive than truth-tellers	Liars are less evasive than truth-tellers
Q2	Guilt	Liars are less emphatic in their denials	Liars are more emphatic in their denials
Q3	Knowledge	Liars are more likely to deny knowledge	No prediction
Q4	Suspicion	Liars are less likely to name another suspect	No prediction
Q5	Vouch	Liars are less likely to name someone who is innocent	Liars are more likely to name someone who is innocent
Q6	Credibility	Liars are less likely to admit that a crime took place	Liars are more likely to admit that a crime took place
Q7	Opportunity	Liars are less likely to admit that they had an opportunity	No prediction
Q8	Attitude	Liars are more likely to voice negative feelings	No prediction
Q9	Think	Liars are more likely to admit to having thought about committing a crime	No prediction
Q10	Motive	Liars are less likely to give a reasonable motive	Liars are more likely to give a reasonable motive
Q11	Punishment	Liars are less likely to suggest serious punishment	Liars are more likely to suggest serious punishment
Q12	Second chance	Liars are more likely to give someone a second chance	Liars are less likely to give someone a second chance
Q13	Objection	liars are more likely to answer in the third person	No prediction
Q14	Results	liars are thought to express less confidence in being exonerated	No prediction
Q15	Loved ones	Liars are less likely to have informed their loved ones	No prediction
Q16	Bait	Liars are thought to express less confidence in being exonerated	No prediction
Non-verbal cues			
Leg crossing		Liars are more likely to	Liars are less likely to
Shifting in chair		Liars are more likely to	Liars are less likely to
Performing grooming behaviours		Liars are more likely to	Liars are less likely to
Leaning forward		Truth-tellers are more likely to	Truth-tellers are less likely to
Establish eye contact		Truth-tellers are more likely to	Truth-tellers are less likely to
Use of illustrators		Truth-tellers are more likely to	Identical to BAI prediction
Answer quickly		Liars are more likely to	Identical to BAI prediction
Sincerity		Liars' answers sound less sincere	Liars' answers sound more sincere
Anxiety-induced behaviours		Liars are more likely to show anxiety-induced behaviours	Liars are less likely to show anxiety-induced behaviours

study, police officers judged the veracity of statements made by lying and truthful suspects during real-life (videotaped) interviews. They found a significant, but negative, correlation between officers reportedly attending to the Inbau et al. cues and accuracy in the lie detection task.

The impression management framework, introduced earlier in the paper to explain liars' tendency to decrease their movements, is also relevant to the verbal responses given in BAI interviews, as it predicts that liars' answers will be more helpful than those of truth-tellers, due to greater efforts to appear convincing. Vrij's (2005a) findings support this assumption. In his experiment, the interviewer challenged the veracity of liars' and truth-tellers' accounts after they narrated what had happened. When they were then asked to repeat what had happened, more truth-tellers than liars refused to do so.

The impression management hypothesis directly opposes Inbau et al. (2001), and hence predicts that (a) liars' answers should be more helpful than those of truth-tellers and (b) liars' non-verbal responses should appear more relaxed than those of truth-tellers. The present experiment focuses on this conflict between the predictions of the impression management hypothesis and those of Inbau et al. These conflicting predictions appear in Table 1.

METHOD

Participants

A total of 40 undergraduate students participated, of whom 21 were male and 19 were female. Their age ranged from 18 to 46 years old and their average age was $M = 21.73$ $(SD = 6.3)$.

Procedure

The experiment took place at a university Students Union. Undergraduates were recruited under the guise of participating in an experiment about "telling a convincing story" for a guarantee £5 and the possibility of earning an additional £10. First, the participants signed an informed consent form, and were allocated randomly to either the truth telling condition or the deception condition. The event used in the truth telling condition is a slightly modified version of Vrij, Akehurst, Bull, & Soukara's (2002). After entering the experimental room, participants allocated to the *truth telling condition* $(N = 20)$ were told by the experimenter that the experiment was run in pairs, and that the other participant was already waiting. They were taken to a room where a confederate (Confederate 1, 'Sam') was sitting alone at a table, posing as another participant. Confederate 1 was played consistently by the same person for all 20 truth-tellers. Both the participant and the confederate were then asked to play a game of Connect 4 and were given some time to get to know each other. Confederate 1 then started to play Connect 4 and chat informally with the participant, steering to one of a few predetermined subject areas. After 1 min, another confederate (Confederate 2) entered the room, apologised for

interrupting and wiped off some mathematical equations from a white board. He/she[4] then left the room after making one of a selection of comments that had previously been scripted. After another minute Confederate 1's mobile phone rang and she excused herself and left the room to answer it, thus leaving the participant alone for a moment. Confederate 1 then returned and Connect 4 play recommenced. After another minute Confederate 3 entered looking for a wallet (the person playing Confederate 3 was also varied from time to time). Confederate 3 also made comments that were based on one of three scripts as s/he looked for the wallet. The wallet had previously been placed in one of several different locations around the room. Confederate 1 (and normally the participant) helped Confederate 3 look for the wallet, which would then be found by either the confederate or the participant depending on its location and Confederate 3 would express relief and leave. A minute later, the experimenter returned with Confederate 3 who claimed that £10 was missing from the wallet. Since Confederate 1 and the participant were the only two people in the room with the wallet they would then be individually interviewed by a police officer. The participant, who was always interviewed first, was then given the following information to read before being taken through to be interviewed by the police officer. 'You are about to be interviewed regarding money that has gone missing from a wallet. When asked about the events that have just recently occurred, explain in as much detail as you can, everything that has just happened in this room. Remember to give as much detail about everything in the room, everyone who has entered and/or exited the room, and everything that has been said and done in the room. The interviewer will then make a judgement, based on what you tell him, about whether or not you are telling the truth. If you manage to convince the interviewer that you did not take the money then you will receive an additional £10 to the £5 you are getting for participating in this study. However, if you do not convince him that you are telling the truth then you will not receive the additional £10, and will have to give a written statement about what actually occurred. You now have a few minutes to familiarise yourself with your story before going into the next room to give your story to the interviewer. Remember to give as much detail as you can in order to make your story convincing'.

The scenario was designed for the purpose of BAI interviewing. For example, Confederate 1 was already in the room before the participant entered, which means that Confederate 1 had the opportunity to have taken the money (see Questions 4 and 7). Confederate 1 left the room for a while, leaving the participant alone in the room. This means that the participant also had the opportunity to take the money (see Questions 4 and 7). Confederate 2 was never alone in the room, hence Confederate 2 could not have taken the money (see Question 5).

Participants in the *deception condition* ($N = 20$) were recruited the same way. They were taken to the same room as the truth-tellers, with the white-board and

---

[4]In order to avoid that all truth-tellers would report the same descriptions, Confederates 2 and 3 were changed from time to time, wearing different clothing. For the same reason, the remarks made by the confederates and the topics discussed during the game of Connect 4 varied. We did not wish all truth-tellers to give descriptions of exactly the same people/clothes etc., because raters would then easily realise whether or not a statement was truthful, just by looking at the similarity between the statements. This is particularly relevant for the free recall part of the interview, which is not discussed in this paper.

equations. They were asked to take £10 out of the wallet. They were then told that their task would be to convince the interviewer that they did not take the money. Instead, they should pretend that they played a game of Connect 4 with a female called Sam. They were given a detailed written account of the event used in the truth-telling scenario, and were given 5 min to familiarise themselves with this alibi and to absorb their story. They were told to give as much detail as possible in order to make their story convincing. Finally, they were told that the interviewer would make a judgement, based on what the participant would tell him, about whether or not the participant was lying. If they managed to convince the interviewer that they did not take the money, they would keep the £10 in addition to the £5 they are getting for participating in this study. However, if they do not convince the interviewer, they would not keep the additional £10, and would have to give a written statement about what actually occurred. The experimenter answered any questions the participant may have and escorted him or her next door for the interview.

The liars did not engage in any of the activities the truth-tellers were engaged in (playing Connect 4, etc.). Instead, the liars took the money out of the wallet, hid it somewhere on their person, and had to pretend that they were engaged in the truth-tellers' activities. They, therefore, had to lie about the entire scenario, including taking £10 from the wallet. The procedure thus reflects a situation where a liar is familiar with the event s/he described but lacks the experience of true participation in that event.

All 40 participants were interviewed by the same uniformed British male police officer. The interviewer was blind to the participant's condition (truth telling or lying). He started the interview by saying "£10 has gone missing from a wallet in the room next door and I have to find out whether or not it was you who took it." After several introductory questions, which gave the participants the opportunity to settle themselves down, the actual interview commenced with the invitation to give a detailed account of what happened when the participant was in the room next door. Although asking such a question is common procedure in BAI interviewing, it is not an actual part of the BAI interview and the replies are therefore not further discussed. After the participants finished their recall, the BAI started. In total 15 questions were asked, and these are all listed in the appendix. Question 15, 'tell loved ones' was omitted as it was not appropriate for this experiment. The bait question, Question 16, was asked although it does not appear to be a standard part of the BAI technique. On average the participants' replies to the 15 BAI questions contained $M = 269.55$ words ($SD = 129.8$) ranging from 94 to 665 words. Veracity had no effect on the number of words spoken, $F(1, 38) = 1.06$, ns.

After the interview, the police officer gave each participant a questionnaire, which he/she completed in another room. The experimenter then told the participants that the police officer had been convinced by their story. Thus, both the truth-tellers and the liars were paid the full £15.[5] The experimenter debriefed participants and answered any questions that they had.

---

[5]This pretence was because it was decided most ethical to pay all participants the same amount, but if participants realised that they would be paid the full amount whether they were convincing or not, then the £10 bonus would no longer be an effective motivation to be convincing.

BAI Coding

All interviews were audiotaped, videotaped and transcribed. The verbal part of the BAI coding was based on the transcribed responses. The BAI questionnaire consisted of 32 items on which truth-tellers and liars should differ according to Inbau et al. (2001). For example, Inbau et al. (2001) report that, compared to truth-tellers, liars are more likely to give naive or evasive replies to the purpose question (Question 1: "What is your understanding of the purpose of this interview?"). Hence, the questionnaire contained two items for the purpose question: "To what extent was the interviewee's answer naive? (i.e., not mentioning the theft)" and "To what extent was the interviewee's answer evasive? (i.e., not mentioning the theft)". Answers could be given on 5-point Likert scales ranging from (1) *not at all* to (5) *totally*. To give another example, Inbau et al. (2001) state that if asked the opportunity question (Question 7: "Who would have had the best opportunity to have taken the money if they had wanted to?") truth-tellers, if they had the opportunity to take the money, would be more open and realistic in disclosing this information, whereas liars would not be likely to point the finger at themselves. Thus, the questionnaire contained the following two items for the opportunity question: "To what extent does the interviewee answer this question open/realistically?" and "To what extent does the interviewee point the finger at him/herself?" Again, answers were given on 5-point Likert scales ranging from (1) *not at all* to (5) *totally*.[6] For the purpose of the analyses several items were recoded so that a higher score always indicated truth telling according to Inbau's assumptions. (In other words, both purpose items in the example mentioned earlier were recoded.) Answers on the items were averaged for each BAI question and each average score could range from 1 to 5. (In other words, the final score on the purpose question was the average score of the two purpose items.) A verbal BAI total score, the summation of the clusters, was compiled excluding the bait cluster. The 'total verbal BAI minus bait' scores could therefore range from 14 to 70. In order to check for interrater reliability a second rater (first author) completed the BAI ratings independently of the first rater (postgraduate research assistant). Table 2 (top) shows the results. High correlations emerged for nearly all 15 questions. The one low correlation (guilt question, Q2) was due to lack of variability: Nearly all participants strongly denied being guilty of the crime. If ratings vary only slightly between raters (the result of low variation in the guilt question) correlations tend to underestimate the true interrater agreement (McGraw & Wong, 1996). Percentage agreement scores were also calculated. The two raters' scores were considered to be in agreement if they differed up to one scale degree from each other. Thus, for each question, if Rater 1's score was 3.0 and Rater 2's score was 3.5, it was considered an agreement. Such extended percentage agreement scores are both acceptable and more often used with 5-point Likert scale ratings (Gödert, Gamer, Rill, & Vossel, 2005). The extended percentage agreement scores (Table 2) suggest strong reliability in BAI coding. This was further supported by the finding that the two total verbal BAI scores correlated highly with each other, $r(40) = .82$. The final verbal BAI scores were the average score of the two coders.

[6]The BAI questionnaire (both verbal and non-verbal parts) is available from the first author.

Table 2. Interrater Reliability Regarding BAI Items

	r (40)	Extended % agreement		r (10)	Extended % agreement
Q1, purpose	0.91	97.5	Crossed legs (R)	0.95	100
Q2, guilt	0.04	100	Shift in chair (R)	1.00	100
Q3, knowledge	0.62	95	Grooming behaviour (R)	–	100
Q4, suspicion	0.90	95	Lean forward	0.67	100
Q5, vouch	0.88	97.5	Direct eye contact	0.89	90
Q6, credibility	0.72	95	Illustrators	–	100
Q7, opportunity	0.94	97.5	Wait before answering	0.87	90
Q8, attitude	0.86	95	Sound sincere	0.75	90
Q9, think	0.89	97.5	Shift posture (R)	1.00	100
Q10, motive	0.50	100	Anxiety-reduced behaviours (R)	0.88	90
Q11, punishment	0.47	82.5	Total non-verbal BAI score	0.94	
Q12, second chance	0.84	92.5			
Q13, objection	0.60	97.5			
Q14, results	0.62	85			
Q16, bait	0.69	95			
Total verbal BAI minus bait score	0.82				

The two raters were blind to the participant's condition (lying or truth telling) and blind to the hypothesis of the study.[7]

The non-verbal part of the BAI coding was based on the videotaped interviews and consisted of 10 items. Coding was carried out by the same two coders who coded the verbal cues. The non-verbal coding referred only to three BAI questions (Questions 2, 3 and 10). Inbau and colleagues suggest that while answering the guilt question (Question 2: "Did you take the money?") the innocent examinee will lean forward in the chair, establish direct eye contact, and may use illustrators to reinforce confidence in the statement, whereas guilty examinees may engage in non-verbal behaviours such as crossing the legs, shifting in the chair, or grooming behaviour. Thus, the BAI coder rated the following six items when the examinee responded to the guilt question: "Does the interviewee cross his/her legs?", "Does the interviewee shift in his/her chair?", "Does the interviewee show grooming behaviour?", "Does the interviewee lean forward in his/her chair?", "Does the interviewee establish direct eye contact?", and "Does the interviewee use illustrators to reinforce the confidence of his/her statement?". Four more items were rated when the interviewee responded to the knowledge and motive questions: "Does the interviewee wait long before answering the question?", "Does the interviewee sound sincere?", "Does the interviewee shift posture in his/her chair?", and "Does the interviewee engage in 'other anxiety-reducing behaviours?". Answers were given on 5-point Likert scales ranging from (1) *not at all* to (5) *totally*. Again, for the purpose of the analyses several items were recoded so that a higher score always indicated truth telling according to Inbau's assumptions.

The coder (and also the second coder, see later discussion) were shown responses only to the three questions being coded, and were able to hear both the

[7]At the time of coding we thought that we were going to test the BAI hypothesis. Note that the codings of both coders correlated strongly despite the first coder being blind to both the BAI and impression management hypotheses.

questions and the interviewee's verbal responses. The latter was necessary to code the vocal aspects (waiting time before answering a question, and sounding sincere). A further reason for using sound was to avoid confusion about which question the coder was coding. Conducting non-verbal coding with sound is common practice in non-verbal communication research (McNeill, 1992; Vrij, 2000). In fact, this is often necessary to accurately discriminate between the different types of arm movements people make (McNeill, 1992; Vrij, in press). As with the verbal component, some items were recoded so that a higher score always indicated truth telling according to Inbau's assumptions. These 10 items were clustered to form the non-verbal component of the BAI scale which could range from 10 to 50. In order to check for reliability, a second coder (the first author) rated 25% of the videotapes ($N = 10$) independently. Correlations for each of the 10 behaviours (see Table 2, bottom) were high (the correlations for two items could not be calculated due to lack of variation). Extended percentage agreement scores further revealed strong consistency between the coders. Also the two non-verbal BAI total scores correlated strongly, .94, suggesting high reliability in the coding. The total BAI score, the score used in the present analyses, is the combination of the verbal and non-verbal BAI score and could range from 24 to 120. We considered it important to look at the total score, rather than just looking at the individual questions, since using the total BAI profile reflects real-life practice (Blair, 1998).[8] Both of the raters always did verbal coding first. We believe it to be highly unlikely that conducting the verbal coding in any way affected the non-verbal coding, because (a) almost 6 months elapsed between the verbal and non-verbal coding, and (b) the coders did not see the entire interview while conducting the non-verbal coding.

Manipulation Checks

Several manipulation checks were carried out. First, participants were asked (i) to what extent they were motivated to appear convincing during the interview, (ii) what they thought the likelihood was of getting the £10, and (iii) what they thought the likelihood was of being made to write a statement. Answers could be given on Likert scales ranging from (1) *very unlikely* to (7) *very likely*. To measure the extent to which participants considered they were inclined to control themselves verbally and non-verbally in order to appear convincing during the interview, 10 statements were used, including: "During the interview I attempted to control my behaviour", "During the interview I attempted to control what I was saying", "During the interview I tried to behave in a way that would appear honest", and "During the interview I tried to say things that would sound honest". Answers could be given on Likert scales ranging from (1) *certainly not* to (7) *certainly*. Those items were clustered into one 'self control' index (Cronbach's alpha = .88).

---

[8]It is unclear from Blair's (1998) writing whether trained BAI interviewers actually compute total scores when applying the BAI method or whether they rely upon their impressions of the suspect's responses. Both methods are likely to be correlated. However, the latter method would be less standardised and therefore less desirable.

Table 3. Manipulation Checks as a Function of Veracity

	Truth		Lie			
	M	SD	M	SD	F(1, 38)	η^2
Motivation	5.70	1.1	5.75	1.0	0.02	
£10 incentive	4.10	1.5	4.20	1.5	0.05	
Writing a statement	3.80	1.6	3.75	1.1	0.01	
Self-control	4.51	1.4	5.27	0.7	4.88∗	.11

∗$p < .05$.

RESULTS

Manipulation checks

Liars and truth-tellers were highly motivated, and there was no difference between the two groups on the motivation item ($M = 5.72$, $SD = 1.0$); see Table 3, The impression management hypothesis does not predict a difference between liars and truth-tellers regarding motivation. It predicts that liars take their credibility less for granted than truth-tellers and therefore try harder to make an honest impression. It does not predict that liars would be more motivated to make an honest impression. In fact, both liars and truth-tellers were promised an equal reward of £15 for being convincing, and it is therefore not surprising that liars and truth-tellers were equally motivated. Only 1 participant (2.5%) reported having not been motivated (scoring lower than 4 on the Likert scale), whereas 35 participants (87.5%) indicated that they were motivated (scoring higher than 4 on the Likert scale).

Liars and truth-tellers did not differ from each other regarding how likely it would be that they would receive a £10 incentive, or having to write a statement. Twelve participants (30%) thought it unlikely that they would receive the £10 incentive and 13 participants (33.5%) thought it likely that they would have to write a statement. In other words, most participants were motivated and the incentive and penalty appeared realistic to them.

In accord with deception literature, liars tried significantly harder to appear honest than truth-tellers. Only six participants (of whom five were truth-tellers) reported that they did not attempt to control themselves in order to make an honest impression (scoring lower than 4 on the Likert scale).

Hypothesis Testing

In order to test the impression management hypothesis an ANOVA was carried out with Veracity (lying vs. truth telling) as the factor and the total BAI score (minus the bait question) as the dependent variable. In line with the impression management hypothesis, but contradicting Inbau's assumptions, liars obtained a higher BAI score than truth-tellers $F(1, 38) = 7.25$, $p < .01$, $\eta^2 = .16$ (see Table 4). We then analysed the individual responses. In the first MANOVA the 15 verbal clusters were included (also the bait question), and in the second MANOVA the 10 non-verbal responses were included. Both MANOVAs revealed significant multivariate effects,

Table 4. BAI Scores as a Function of Veracity

	Truth		Lie			
	M	*SD*	*M*	*SD*	*F*(1, 38)	η^2
BAI total score − bait	80.49	5.3	84.79	4.8	7.25**	.16
Q1, purpose	3.56	1.4	4.43	1.0	4.81∗	.11
Q2, guilt	4.98	0.1	5.00	0.0	3.35	
Q3, knowledge	1.11	0.2	1.39	0.5	4.11	
Q4, suspicion	3.25	1.4	2.55	1.7	2.10	
Q5, vouch	2.68	1.7	3.72	1.5	4.26∗	.10
Q6, credibility	3.30	0.9	2.84	0.9	2.66	
Q7, opportunity	3.06	1.1	3.53	1.2	1.63	
Q8, attitude	4.25	1.2	3.88	1.1	1.05	
Q9, think	3.83	1.5	4.38	0.9	1.93	
Q10, motive	3.30	0.3	3.38	0.3	0.77	
Q11, punishment	3.28	0.8	3.39	0.7	0.25	
Q12, second chance	3.72	1.0	4.22	1.0	2.57	
Q13, objection	3.78	0.6	3.54	0.5	1.99	
Q14, results	3.71	0.8	3.92	0.7	0.83	
Q16, bait	4.34	0.7	3.81	0.6	6.54∗	.15
Crossed legs (R)	4.15	1.0	4.80	0.6	6.24∗	.14
Shift in chair (R)	4.25	1.0	4.65	0.6	2.31	
Grooming behaviour (R)	4.95	0.3	4.95	0.3	0.00	
Lean forward	1.35	0.7	1.30	0.7	0.06	
Direct eye contact	4.05	0.9	4.30	0.9	0.82	
Illustrators	1.20	0.7	1.10	0.4	0.29	
Wait before answering	1.25	0.6	1.05	0.2	2.27	
Sound sincere	4.15	0.7	3.95	0.7		
Shift posture (R)	3.85	1.3	4.65	0.7	5.92∗	.14
Anxiety-reduced behaviour (R)	3.45	1.5	3.85	1.2	0.88	
BAI total score + bait	84.83	5.4	88.60	4.6	5.57∗	.13

Note. Some scores were recoded so that a higher score always indicates truth telling according to Inbau et al.'s assumptions.
*p<.05. **p<.01.

$F(15, 24) = 2.29, p < .05, \eta^2 = .59$ for the verbal clusters and $F(10, 29) = 2.52, p < .05,$ $\eta^2 = .46$ for the non-verbal responses.

The results of the univariate tests are presented in Table 4. Table 4, reveals that five significant differences emerged. Inbau's assumption that truth-tellers will obtain higher scores than liars occurred for only one significant finding, the bait question. As can be seen in Table 4, truth-tellers expressed more confidence in being exonerated than liars. The bait question does not form part of the official BAI interview. Other significant findings contradicted Inbau's assumptions. Thus, truth-tellers were more naive and evasive while answering the purpose questions, and were less likely to name someone who they felt certain did not take the money (vouch question). Truth-tellers were also more likely to cross their legs and were more likely to shift posture (both items were recoded, see Method section) than liars.

Since the bait question, the only question to have had an effect predicted by Inbau and colleagues, was not included in the total BAI score (because it is not part of the standard BAI procedure), a new BAI total score was compiled consisting of the old BAI score with the addition of the bait question. The new BAI score showed the same pattern as the old BAI score: Liars had higher scores than truth-tellers, in

agreement with the impression management hypothesis, but contradicting the BAI assumptions.

DISCUSSION

The present experiment is the first empirical test of the BAI where the ground truth has been established. BAI is an interview technique developed by Inbau et al. (2001) and designed to evoke different verbal and non-verbal responses in liars and truth-tellers. Indeed, the BAI questioning led to differences between liars and truth-tellers but the difference was in the opposite direction to that anticipated by Inbau et al. They expected liars to be less helpful in investigations and to exhibit more nervous behaviour. In fact, liars were more helpful and displayed less nervous behaviour. For example, liars crossed their legs less and shifted posture less than truth-tellers. That is, liars sat more still and made fewer movements than truth-tellers, which is in alignment with deception research (DePaulo et al., 2003). The deception literature also provides an explanation for why liars were more helpful and displayed less nervous behaviour. Liars are typically more concerned with impression management than truth-tellers. Liars take their credibility less for granted than truth-tellers and are more aware of their responses and of the impact these responses have on others. Perhaps the main problem with Inbau's approach is that it fails to take into account the importance of impression management.

In order to challenge our findings, one could argue that the present experiment was a laboratory study rather than an examination of real life police interviews with suspects, and that Inbau's assumptions are valid in real life suspect interviews. We find it hard to believe that this would be the case. We do not deny that there are differences between real-life interviews and our experimental situation. For example, the positive consequences of being believed (£10 incentive) and negative consequences of being disbelieved (threat of writing a report) in our experiment were considerably lower than the stakes in real life police interviews. One would expect that a low-stakes study would result in an absence of significant differences between liars and truth-tellers, and, if that had been the outcome of the present study, we could not have legitimately challenged Inbau's assumptions. However, we did find significant differences between liars and truth-tellers but they were *opposite* to that predicted by Inbau et al. In other words, if we still hold on to Inbau's assumptions, we need to explain why liars and truth-tellers would respond in a qualitatively different way rather than a quantitatively different way in lower stakes situations compared to higher stakes situations. We do not know how to theoretically explain such a switch in responses. Also, as noted earlier, the findings of the present experiment are in alignment with the deception literature and deception theory (DePaulo et al., 2003).

We do applaud the effort that has gone into BAI to introduce an interview protocol designed to evoke different verbal and non-verbal responses in liars and truth-tellers. As we mentioned in the beginning of the paper, the question of how interview styles might benefit non-verbal lie detection has been ignored to date. We

believe that such an approach has potential but it is essential that such interview styles be grounded in sound theory. For example, one reason why cues to deceit emerge is because liars experience more cognitive load than truth-tellers (DePaulo et al., 2003), and research suggests that this cognitive load aspect might be particularly important in police suspects (Mann, Vrij, & Bull, 2002). Police interviewers could use this by employing interview techniques that increase the suspect's cognitive demand. This should have a greater effect on liars than on truth-tellers, thus facilitating discrimination between them. There are several ways in which cognitive demand could be increased and we will give two examples (see Vrij, 2004, and Vrij & Mann, 2005, for more on this topic). First, interviewers can request suspects to elaborate on issues they have previously mentioned. This might be more difficult for liars than for truth-tellers, especially if they did not expect to be asked for such elaboration and, hence, they may not have prepared responses to such questions prior to the interview. Second, it is often found in the deception research that truth-tellers tend to tell their stories in a more *un*structured way than liars (Vrij, 2000, 2005b). That is, liars tend to tell their stories in a more fixed chronological order (this happened first, and then this, and then that, and so on) than truth-tellers. It has been suggested that it is difficult for liars to fabricate a story in a non-chronological order (Steller, 1989; Köhnken, 1999; Zaparniuk, Yuille, & Taylor, 1995). Lie detectors could exploit this difficulty by asking interviewees to tell their stories in a non-chronological order, for example in reverse, order. Presumably, this novel output order should be relatively easy for truth-tellers to generate, since they are not under as much cognitive load, whereas it may be difficult for liars who are burdened by a more severe cognitive load. Whether these interview styles actually work and will benefit lie detectors is currently unknown. We believe that the efficiency of such interview styles should be empirically tested before they are put into practice.

ACKNOWLEDGEMENT

Appendix: The Behaviour Analysis Interview Questions in Case of Alleged Theft of Money

Q1, *purpose*. What is your understanding of the purpose of this interview?
Q2, *guilt*. Did you take the money?
Q3, *knowledge*. Do you know who did take the money?
Q4, *suspicion*. Who do you suspect might have taken the money?
Q5, *vouch*. Is there anyone other than yourself who you feel certain did not take the money?
Q6, *credibility*. Do you think that someone did actually purposefully take the money?

Q7, *opportunity*. Who would have had the best opportunity to have taken the money if they had wanted to?

Q8, *attitude*. How do you feel about being interviewed about the missing money?

Q9, *think*. Have you ever just thought about doing something like taking the money from a wallet that is just lying around?

Q10, *motive*. Why do you think someone did take the money?

Q11, *punishment*. And what do you think should happen to the person who took the money?

Q12, *second chance*. Do you think there are any circumstances under which the person who took the money should be given a reprieve?

Q13, *objection*. Tell me what would you stop from taking the money?

Q14, *results*. Once we completed our investigation, what do you think the results will be in respect to your involvement in taking the money?

Q15, *tell loved ones*. Who did you tell about your interview with me today?[5]

Q16, *bait*. There is a CCTV system working in that room and we can get access to the tapes. If we look at those tapes, is there any reason that we will see you taking £10 out of the wallet? I am not saying that we will see you taking the money, I just want to know your views about what the likelihood of this might be?[5]

Derived from Inbau et al. (2001).

REFERENCES

Blair, J. P. (1998). Detecting deception: The effects of Reid Behavioral Analysis Interview Training. Unpublished thesis, Western Illinois University.

DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 51–70). Dordrecht, The Netherlands: Kluwer.

DePaulo, B. M., LeMay, C. S., & Epstein, J.A. (1991). Effects of importance of success and expectations for success on effectiveness at deceiving. *Personality and Social Psychology Bulletin*, *17*, 14–24.

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

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, politics and marriage*. New York: W. W. Norton & Company.

Ekman, P. (1997). Deception, lying, and demeanor. In D. F. Halpern & A. E. Voiskounsky (Eds.), *States of mind: American and post-Soviet perspectives on contemporary issues in psychology* (pp. 93–105). New York, NJ: Oxford University Press.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication*, *22*, 353–374.

Frank, M. G. & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. *Journal of Personality and Social Psychology*, *72*, 1429–1439.

Frank, M. G., & Feeley, T. H. (2003). To catch a liar: Challenges for research in lie detection. *Journal of Applied Communication Research*, *31*, 58–75.

Gödert, H. W., Gamer, M., Rill, H. G., & Vossel, G. (2005). Statement Validity Assessment: Interrater reliability of Criteria-Based Content Analysis in the mock-crime paradigm. *Legal and Criminological Psychology*, *10*, 225–245.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. Chichester: Wiley.

Horvath, F., Jayne, B., & Buckley, J. (1994). Differentiation of truthful and deceptive criminal suspects in behaviour analysis interviews. *Journal of Forensic Sciences*, *39*, 793–807.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogation and confessions* (4th ed.). Gaithersburg, Maryland: Aspen Publishers.

Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist*, *52*, 221–233.

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.

Kassin, S. M., & Norwick, R. J. (2004). Why people waive their Miranda rights: The power of innocence. *Law and Human Behavior*, *28*, 211–221.

Kleiner, M. (Ed.). (2002). *Handbook of polygraph testing*. San Diego, CA: Academic Press.

Köhnken, G. (1999, July). Statement Validity Assessment. Paper presented at the pre-conference program of applied courses 'Assessing credibility' organised by the European Association of Psychology and Law, Dublin, Ireland.

Köhnken, G., & Steller, M. (1988). The evaluation of the credibility of child witness statements in German procedural system. In G. Davies & J. Drinkwater (Eds.), *The child witness: Do the courts abuse children? (Issues in Criminological and Legal Psychology, no. 13)* (pp. 37–45). Leicester, United Kingdom: British Psychological Society.

Mann, S., Vrij, A., & Bull, R. (2002). Suspects, lies and videotape: An analysis of authentic high-stakes liars. *Law and Human Behavior*, *26*, 365–376.

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

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*, 30–46.

McNeill, D. (1992). *Hand and mind*. Chicago: The University of Chicago Press.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, *26*, 469–480.

Steller, M. (1989). Recent developments in statement analysis. In J. C. Yuille (1989). *Credibility assessment* (pp. 135–154). Deventer, The Netherlands: Kluwer.

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, NJ: Springer-Verlag.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and its implications for professional practice*. Chichester: Wiley.

Vrij, A. (2004). Invited article: Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology*, *9*, 159–181.

Vrij, A. (2005a). Cooperation of liars and truth-tellers. *Applied Cognitive Psychology*, *19*, 39–50.

Vrij, A. (2005b). Criteria-Based Content Analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, *11*, 3–41.

Vrij, A. (in press). Non-verbal communication and deception. In V. Manusov & M. L. Paterson (Eds.), *The handbook of non-verbal communication*. Thousand Oaks: Sage.

Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2002). Will the truth come out? The effect of deception, age, status, coaching, and social skills on CBCA scores. *Law and Human Behavior*, 26, 261–283.

Vrij, A., & Mann, S. (2005). Police use of non-verbal behavior as indicators of deception. In R. E. Riggio & R. S. Feldman (Eds.), *Applications of non-verbal communication* (pp. 63–94). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Zaparniuk, J., Yuille, J. C., & Taylor, S. (1995). Assessing the credibility of true and false statements. *International Journal of Law and Psychiatry*, *18*, 343–352.