

The behavioural analysis interview: clarifying the practice, theory and understanding of its use and effectiveness

Frank Horvath[†], J. P. Blair[‡] and Joseph P. Buckley[§]

[†]Michigan State University, 108 Columbia Club Drive-W, Blythewood, SC 29016, USA.

Tel: (803) 751 9196; Fax: (803) 751 9178; email: Horvath@msu.edu

[‡](Corresponding author) Texas State University, Department of Criminal Justice, 601 University Drive, San Marcos, TX 78666, USA. Email: pete.blair@txstate.edu or jpbclair1@yahoo.com

[§]John E. Reid and Associates, Inc., 209 W Jackson Blvd., Suite 400, Chicago, IL 60606, USA. Tel: 312 583 0700; Fax: 312 583 0701; email: info@reid.com

Received 23 January 2007; accepted 3 April 2007

Keywords: police interviewing, police interrogation, behavioural analysis interviewing, the BAI

Frank Horvath is professor emeritus, School of Criminal Justice, Michigan State University.

J. P. Blair is an assistant professor of criminal justice at Texas State University.

Joseph P. Buckley is the president of John E. Reid and Associates, Inc.

ABSTRACT

The Behavioural Analysis Interview (BAI) is the only questioning method that has been developed specifically to help investigators sort those who are likely to be 'guilty' from those who are not. In its typical application the BAI is a pre-interrogation interview that is used to focus interrogational effort; however, it also can be used independently in order to circumscribe investigative efforts in those cases in which there is a fixed and relatively large number of 'suspects'. In this paper an overview of the BAI process is provided and the findings and limitations of the extant bodies of field and laboratory research on the BAI are discussed. The paper concludes with suggestions to guide future research on the BAI.

INTRODUCTION

Everyone lies to the police (Ericson, 1981; Innes, 2003; Simon, 1991). Guilty suspects lie for obvious reasons. Witnesses lie out of fear of retribution; sometimes to cover for perceived personal failures. Victims lie to conceal their true involvement in matters that precipitated a crime. Even innocent suspects lie, often because they perceive no need to cooperate with the police. And, of course, persons in all of these roles also tell the 'truth'.

What and whom the police find credible are, at the core, the essence of the questions addressed in the police investigative task (Ericson, 1981; Horvath & Meesig, 1996; Innes, 2003; Simon, 1991). This is a dilemma. How do the police sort the truth-tellers from the liars; the credible story from the spurious? From the police point of view, this is a learned, on-the-job task, one in which they, if observant, ambitious and skilful, believe they eventually will become adept. From the perspective of a trained scientific observer, however, discerning

between a 'lie' and a non-lie, even for those whose professional interest focuses on the problem, cannot be reduced to a simple, formulaic task. In fact, it is, according to some reports, as difficult for most 'professionals' as for others to determine when a lie is told (Bond & DePaulo, 2006; Ekman, O'Sullivan, & Frank, 1999). There is, however, some evidence that shows that, when 'high stakes' lies are assessed, police officers experienced in interviewing are able to detect truths and lies above chance levels, with an accuracy rate of about 65 per cent (Mann, Vrij, & Bull, 2004). Unfortunately, that study did not compare police officers with non-professionals; consequently, it did not show that 'professional' persons are better at lie detection than others. The accuracy statistic, however, suggests the validity of what seems to be an accepted truism, in the scientific literature at least, that detecting lies is quite difficult and not done, even in the best of circumstances, with a high degree of accuracy (Bond & DePaulo, 2006; DePaulo et al., 2003).

THE BAI

One approach that police officers can use to help them distinguish between those who are 'guilty' and those who are 'innocent' is called the Behavioural Analysis Interview (BAI). In fact, this is the only questioning method that has been developed specifically for this purpose. In its typical application the BAI is an interview process that the police use prior to a formal, accusatory interrogation. The purpose of the BAI is to help investigators sort those who are likely to be guilty from those who are not and thus to interrogate only those in the former category.

According to an article by Blair and Kooi (2004), training in the use of the BAI has been actively promoted as a part of a training protocol in the application of the Reid

Technique, an approach to interrogation that may be the most well-known method in the world. Blair and Kooi maintain that over 150,000 police personnel have been trained in the use of the BAI throughout the world. But, they also point out that the basis for the BAI has not been solidly established in the scientific literature. This suggests that the BAI may lead the police to interrogate (using accusatory questioning) a truly innocent person and to fail to take appropriate action with respect to an actually guilty suspect. There is, therefore, a need for both practitioners and academics to engage in further research on the BAI and the Reid Technique in order that the police may rely on these processes with appropriate confidence.

The call for additional research on the BAI by Blair and Kooi (2004) would seem to be addressed in a recent study by Vrij, Mann, and Fisher (2006). They reported what is the first experimental, laboratory-based evaluation of the BAI. However, their report, coupled with other information found in the literature, demonstrated significant misunderstandings of the development, structure and research regarding the BAI. This article is intended to present an overview of the BAI and the research which supports it so that practitioners and researchers will not be misled by findings that have little, if any, ecological validity. In this paper a brief overview of the BAI process is offered first. Then the strengths and limitations of the current body of BAI research are discussed. Finally, suggestions are made about how to improve research on the BAI in the future.

Before getting to the heart of this paper, the key components and assumptions of the BAI will be described. More detail can be found in Blair (1999), Horvath (1973), Horvath and Jayne (1990), Horvath, Jayne, and Buckley (1994) and Inbau, Reid, Buckley, and Jayne (2001).

OVERVIEW OF THE BAI

General issues

A few special observations are in order at the outset. First, it is to be noted that it is common to find that the term 'guilt' is used synonymously with 'deception' and 'lying' in some contexts; whereas truthfulness is used to indicate 'innocence', 'honesty', 'truth-telling' and so forth. The difference between these terms is not always clear, nor is it evident from the context in which they are used precisely what is being referred to. For example, Vrij et al. (2006) stated that the BAI 'is designed to evoke behavioural responses to detect lying . . .' (p. 330). However, the BAI is not used to detect 'lying', at least not in the sense that there is an attempt to determine if a particular statement or message from a 'sender' is or is not true (as is commonly done in 'deception' research).

Rather, the BAI is used to detect 'liars', persons who deliberately withhold critical information regarding their knowledge of or involvement in a matter under investigation. (Maybe 'withholders' is a better term.) During the BAI a liar may, and usually does, tell a 'lie'. For example, she may say in response to a direct question, 'No, I did not steal that \$500' when, in fact, she did. But the detection of that 'lie' is not critical to the outcome of the BAI. It is only the completed BAI that leads to the inference that the 'liar' is indeed withholding relevant information and thus warrants additional investigative attention. In short, the BAI is perhaps best seen in the light of a 'guilty person' as opposed to a 'deception detection' paradigm as these terms have been used customarily in the 'lie detection' literature.

A properly conducted BAI leads to a decision of either 'Eliminated from suspicion', or 'Not eliminated from suspicion'. Outcomes in the latter category might result in follow-up questioning or the use of

additional targeted, investigative resources. For this reason the terms 'guilty' and 'innocent' are used here to refer, respectively, to those who are not and those who are eliminated from suspicion.

A second point to be made clear is that the BAI is not exclusively a 'police' questioning approach, contrary to what some have suggested (Vrij et al., 2006). Although there is certainly widespread use of the BAI in policing, it is also the case that the BAI is used in many other environments. The major and most realistic assessment of the BAI carried out to date, by Horvath et al. (1994), did not involve police officers, police interviewers or 'suspects' who were in police custody. All interviewers were private employees of John E. Reid and Associates and all 'suspects' were employees of commercial firms suspected of involvement in independent specific thefts of money or property. All questioning took place on the premises of a private organisation without any connection to a police agency. The effectiveness of the BAI in a police environment has never been formally studied.

Rationale of the BAI

It is believed that it is the 'guilty' person's underlying consciousness of involvement in the offence under investigation that is the foundation for the differential attitudes and behaviours of 'innocent' and 'guilty' persons during the BAI. Perhaps the use of a thought experiment will illustrate this. Imagine that you are at work as a bank teller and one day you are told that money is missing. You know you did not take the money. What are the first thoughts that enter your mind? It is likely that they will be 'I wonder what happened?'. Your initial thought might be that the money was misplaced. You do not want to believe that someone, a colleague, stole it. As your self-inquiry proceeds, however, and you learn that the money was not simply misplaced,

you turn to 'what happened' and 'whodunit'. You replay the events of the day in your head. You also consider that since you are certain you are not responsible, someone you work with might be. These considerations will be based not only on your observations of co-workers' activities (ie opportunity and access), but also your personal knowledge of their personality and the circumstances of their lives (ie propensity and motive).

This response to such an event, the 'Sherlock Holmes Effect', is quite common. When an incident such as this occurs there is a tendency for involved persons to try to solve it. Observations over time show that innocent persons, though initially reluctant, are willing to discuss such thoughts (Horvath, 1973; Reid & Arther, 1953; Reid & Inbau, 1977). With prompting, they often reveal their suspicions of those whom they believe are most likely to be guilty.

Guilty persons on the other hand respond differently, and for good reason. They cannot share as much relevant information because that would ultimately lead to their detection, something they want to complicate, not simplify. Additionally, guilty persons do not experience the Sherlock Holmes Effect because they already know 'whodunit' and how it was done. Instead, guilty persons are consumed with thoughts of how to conceal their crime and how to misrepresent information that implicates them.

In the context of the established deception theories, the BAI can be viewed as a special case of the self-presentation theory presented by DePaulo et al. (2003). Both the innocent and the guilty try to craft a self-presentation that will make them appear as an innocent person. However, the guilty person's presentation will usually be less compelling, partly because of misperceptions regarding how truly innocent persons actually behave, and partly because the guilty person cannot share as much

information as the innocent person, for to do so would lead to detection. The inability of the guilty person to share information typically leads to verbal responses to interviewer questions that are shorter and tend to indicate that the guilty person has not given a great deal of thought as to what happened and 'whodunit'; evidence of the Sherlock Holmes Effect is unlikely. The truthful person's responses tend to be longer, more detailed, and show that the person has given appropriate thought to the situation at hand.

In order for the Sherlock Holmes Effect to be observed, several assumptions seem to be necessary. The first is that some event of importance to the people who are involved must have occurred. This assumption is generally met whenever a criminal act occurs in a person's workplace or immediate neighbourhood. The second is that the involved persons must be aware that the event did occur or is likely to have happened. Again, this assumption is easily met. In the workplace, people talk about shortages and the staff often engage in extensive attempts to locate missing money or objects before they are considered stolen. The third assumption is that the involved persons must have time to consider what happened, how it happened and who might be responsible. This is generally not an issue in police or corporate investigations because investigative interviews generally take place hours or even days after the event occurs. This time provides an opportunity not only for the involved persons to think about the event themselves, but also for them to gather information about the event from other people. Fourth, the involved persons must have some knowledge about others who are involved and about the environment in which the event occurred. In a corporate environment, the employees know the other people in their immediate working environment and they typically have a general idea of what is going on in

each other's lives. They also know the procedures and the day-to-day occurrences at their workplace. In a criminal case, the residents of a neighbourhood may or may not know their neighbours and the general events that are occurring in each other's lives. The extent of this knowledge will depend on how tightly knit the neighbourhood is. The more closely knit the neighbourhood the more likely this condition will be met. Finally, the innocent person must have an interest in helping the investigation to succeed. In most cases this is due either to a desire to see justice done or to be exonerated of suspicion. However, in some situations it is possible that the innocent person has no interest in helping the investigation. For example, if someone close to the innocent person is known (by the person) to have committed the crime in question, the helpfulness typically observed may not be present. If these five assumptions are not met, the Sherlock Holmes Effect would be less likely to be experienced and therefore, the utility of the BAI would be more limited.

Viewed in this context, it is again important to point out that the BAI is not a procedure designed to evaluate directly at any particular point whether a person is lying or telling the truth. Rather, the BAI is used to assess the likelihood of involvement in a matter under investigation, often partially revealed by an attitude toward the investigation and the interview at hand. Field experience and field-derived data show that those who are not involved are more likely to be helpful, cooperative and confident; they show honest concern about the investigation and reveal sincerity and spontaneity in their responses during the BAI (Horvath, 1973; Reid & Arther, 1953; Reid & Inbau, 1977). Those who are involved in the matter at issue show other behaviours; they tend to be guarded and defensive; they are generally less helpful and sometimes apathetic. These characteristics

are revealed throughout the entire BAI, not just during a response to a particular question or set of questions, and they influence the message content that is evaluated during the BAI. However, it is important to note that because the research on the BAI to date shows a greater tendency for 'guilty' persons to be judged as 'innocent' than for the opposite outcome, it is possible that the BAI, even in the best of circumstances, may not be effective with persons who are sufficiently skilled at 'deceiving' others (Horvath et al., 1994).

Conducting the BAI

The BAI is typically conducted in a private setting where the interviewer sits directly in front of the suspect at a distance of 4.5 to 5 feet. The interviewer's demeanour is non-judgmental and non-accusatory throughout the interview; this is so even when the interviewer believes that the suspect may have lied to a particular question or at a specific point in the BAI. It is believed that if this non-judgmental and non-accusatory demeanour is not maintained, accurate assessments of guilt (deception) or innocence (truth) are less likely because the behaviour of the interviewer will lead the interviewee to distrust or become resentful of the interviewer and perhaps to manifest misleading behavioural indicators (Inbau et al., 2001). In addition to this concern, misleading behaviours could also result from the tendency of some interviewees to 'mirror' certain features of an interviewer's behaviour (Chartrand & Bargh, 1999). While those who practise the BAI are aware of this and other possibilities that might distort accurate observations, there is no research of their effect, if any, on BAI outcomes.

During the course of the interview, the interviewer takes careful written notes documenting the questions asked of the suspect, as well as the suspect's verbal responses and non-verbal behaviours.

Because of this, there are short periods of silence separating each question. Of course, if an interviewee were to decide not to respond to the interviewer's questions during the BAI, or to respond with only a 'no comment' stance, the effectiveness of an assessment of behaviours would be greatly reduced, perhaps rendered valueless.

The first several minutes of the interview are spent obtaining background information from the suspect. This information, of course, establishes personal demographic data of interest to the interviewer. In addition, however, the collection of such relatively neutral data permits the interviewer to evaluate and note the suspect's 'normative' behaviour, in particular, eye contact, response timing, spontaneity and general nervous tension. During the remainder of the interview, the suspect is asked two different categories of questions, generally in separate time periods. These are investigative and behaviour-provoking questions. Investigative questions concern such things as the suspect's actions, opportunity, access, motivations and propensity to commit the crime. It is important to note that these questions are as important to the BAI process as are the behaviour-provoking questions. For example, consider a case wherein a suspect is seen on video entering the area from which money was stolen. The suspect denies being in that area during the asking of the investigative questions. In such an event, it is highly unlikely that a suspect would be eliminated from suspicion based solely on his or her responses to the behaviour-provoking questions. The normative, investigative and behaviour-provoking questions, in other words, are used in a complementary way.

The behaviour-provoking questions are specifically asked to elicit differential verbal and non-verbal behaviours from 'innocent' and 'guilty' suspects. Over the last 45 years many different behaviour-provoking questions have been developed. Verbal responses

typically given by 'innocent' and 'guilty' persons to some of these were originally documented by Horvath (1973); they have been further developed based on empirical observations and theoretical expectations (Jayne & Buckley, 1999)

Specific guidelines govern the interpretation of messages and behaviours elicited during a BAI (Inbau, Reid, & Buckley, 1986; Inbau et al., 2001; Jayne & Buckley, 1999). These are designed to protect against errors, especially false positive outcomes (reporting an innocent suspect as guilty). These guidelines include the following:

1. Evaluate deviations from the suspect's normal behaviour. There are no unique behavioural indicators which are always associated with 'innocence' or 'guilt'. It is the suspect's deviation from his or her 'normal' behaviour which is significant. Such 'normal' behaviour is reflected especially in response to the questions regarding personal history, which are generally less emotionally provocative for 'guilty' and 'innocent' persons. For the 'innocent' person the items in the investigative and behaviour-provoking categories are less provocative than for the 'guilty' person. Therefore, it is the consistency within and between categories that is important.
2. Evaluate the non-verbal behaviours as they occur in response to a particular question and be attentive to the repetition of the behaviours, that is, their consistency across questions.
3. Evaluate verbal and non-verbal behaviours in conjunction with each other. Look for discrepancies between the two channels.
4. Consider underlying factors which could affect the validity of behaviour analysis. Examples of these include the seriousness of the offence, what the suspect was told about the offence, the

suspect's emotional stability, and the suspect's cultural and social environment.

5. Consider the suspect's behaviour in conjunction with factual analysis. When there is a discrepancy between known facts and a suspect's behavioural indicators, the interviewer must be cautious in rendering a definite opinion of the suspect's involvement.

Examples of behaviour-provoking questions

There are many behaviour-provoking questions that may be asked during a BAI investigating an employee theft. Several examples of these, each dealing with a different aspect of the array of questions included in a BAI, are provided here to illustrate how such questions are used and how responses are assessed. More complete descriptions of these and other questions are found in Horvath (1973), Horvath et al. (1994), Inbau et al. (1986), Inbau et al. (2001).

In these three examples, the employee, Andy, is suspected of stealing a deposit from a department store. The interviewer's questions are preceded by the letter 'I' and are followed in parentheses with the abbreviated word (in italics) used to denote the question; the suspect's responses are indicated by an 'S'. Two responses accompany each behaviour-provoking question; the first response is typical of a suspect not involved in the matter ('innocent') while the second is more indicative of a 'guilty' suspect (see Horvath, 1973; Inbau et al., 1986). It should be noted that during an actual BAI if the suspect's verbal response or the non-verbal behaviours are ambiguous, the interviewer may ask a follow-up question in an attempt to clarify the suspect's status.

- I: (*Purpose*) Andy, what is your understanding of the purpose of this interview with me today?

S: (*Innocent*) Well, last Sunday morning when the bookkeeping department was counting up the deposits they found that the \$3,200 deposit from the men's department was missing and I know that I put it in the safe. So the reason I am here is to prove that I didn't steal it.

S: (*Guilty*) Well, I guess they may have misplaced a deposit envelope and I'm just here to help them find out what might have happened to it.

I: (*Attitude*) How do you feel about being interviewed concerning this missing deposit?

S: (*Innocent*) Oh, I don't mind at all. I want to prove to them that I didn't steal it and hopefully through these interviews they will be able to catch the thief.

S: (*Guilty*) I don't feel one way or the other. It's something that I have to do to keep my job.

I: (*Bait*) Andy, something you may not be aware of is that most drop safes have a counting mechanism on the underside of the drop slot. Very simply the force of an envelope entering the safe causes the counter to advance in increments of one. Now if you in fact did put that envelope in the safe last Saturday the counter should read 11, because that's the total number of envelopes that should have been dropped. Now I don't know if this particular safe is equipped with that mechanism, but if it is, can you think of any reason why the counter would indicate 10 drops instead of 11?

S: (*Innocent*) If it does it's not from my envelope because I know for sure I put that envelope in the safe.

S: (*Guilty*) Gee . . . , I don't know very much about mechanical things, maybe it got stuck or something.

Non-verbal behaviours

Many different non-verbal behaviours are also evaluated during the BAI. While Inbau

et al. (2001) often use specific behaviours as examples of a guilty or innocent response, it must be remembered that it is the consistency of and the change from a person's normative behaviour, not the behaviour itself, that is evaluated. Commonly assessed vocalic behaviours include rate of speech, response latency and length of response. Commonly assessed kinesic behaviours include smiles, head movements, posture changes, illustrators, adaptors, foot and leg movements and eye contact. It must also be remembered that these behaviours are evaluated in conjunction with the suspect's verbal statements.

RESEARCH ON THE BAI

Field studies

Historically, the foundation for the BAI was the early report by Reid and Arther (1953) regarding the behaviour of persons undergoing polygraph examinations. Their report provided the grounding for an empirical evaluation of polygraph subjects' behaviours by Horvath (1973). His research on the 'structured pre-test interview', the portion of a polygraph examination prior to the collection of physiological data, triggered the development of what is known as the BAI, a non-instrumental interviewing method used to assess the likelihood of a person's involvement in a specific matter under investigation.

Following the early Horvath (1973) evaluation, a second, larger scale and more carefully controlled assessment of the BAI was carried out (Horvath et al., 1994). In this study, four evaluators, all highly trained and experienced in the analysis of behavioural information using the BAI, made independent judgments of the innocence or guilt of real-life suspects who were undergoing BAI interviews. These interviews had been audio-visually recorded and

were judged by the evaluators in four conditions:

1. A 'Written' condition in which they had access only to transcribed responses to the protocol of behaviour-provoking questions asked of the suspects.
2. An 'Out-of-context, without audio' condition in which evaluators observed a video-taped replay of the suspects' non-verbal behaviour displayed when responding to each of the behaviour-provoking interview questions; evaluators were unable to hear the suspects' verbal responses but were aware of the question to which the suspect was responding (eg 'Purpose', 'Attitude', 'Bait', etc).
3. An 'Out-of-context, with audio' condition in which the suspect's verbal and non-verbal behaviours were observable.
4. An 'In-context' condition in which each suspect's responses to the protocol of interview questions was observable in a sequence similar to that in the original BAI.

In the first three conditions evaluators made decisions about the suspects' 'innocence' and 'guilt', expressed the degree of confidence in those decisions, and rated certain behavioural and attitudinal dimensions of the suspects as they independently observed 786 segments of written or AV-taped questions. In the In-context condition, evaluators made two separate decisions of 'guilt' and, for each, expressed their degree of confidence. The first decision was made by evaluators as they judged separately each suspect's response to each question asked of each suspect; the question protocol was presented in the same, fixed sequence for each suspect. The second decision, a 'composite' judgment, was made after each evaluator had viewed the entire protocol of questions asked of each suspect. These 'composite' judgments were rendered by

evaluators after they had observed each suspect in a context similar to the actual BAI from which the AV recording had been derived.

Evaluators' accuracy, calculated by combining the four evaluators' decisions of 'guilt' into a single-Group decision, showed that when inconclusive judgments were excluded, correct decisions on innocent suspects averaged 80 per cent, 79 per cent, and 88 per cent in the Written, No-audio, and With-audio conditions respectively. Evaluators' In-context decisions were 88 per cent correct when they were made on observations of separate interview questions. On guilty suspects, evaluators' judgments, excluding that of 'inconclusive', were correct 53 per cent in the Written condition and 75 per cent, and 76 per cent in the No-audio and With-audio conditions, respectively. In the In-context condition when questions were scored separately, decisions were 64 per cent correct.

The different conditions of evaluation had a significant effect on evaluators' overall accuracy. On innocent and guilty suspects the lowest accuracy and the highest rate of inconclusive outcomes was obtained when only non-verbal behaviour was observed (the Out-of-Context, without-audio condition). When evaluators were able to observe both verbal and non-verbal behaviour in a context which closely resembled the real-life circumstance (In-context) and in which they rendered a 'composite' decision, they obtained the highest overall accuracy; excluding inconclusive judgments, they were correct in 86 per cent of their 'composite' judgments on innocent suspects and 83 per cent on guilty suspects. (Inconclusive judgments varied, of course, from evaluator to evaluator and across the different observational methods. To summarise, however, the average percentage of inconclusive judgments in the In-context mode, composite judgments, on 'Truthful'

persons was 14 per cent; on 'Deceptive' persons, 17 per cent.)

It is of interest to note here that analysis of objective scoring of the behaviours of the suspects in the Horvath et al. (1994) study, revealed significant differences between innocent and guilty suspects with respect to verbal and non-verbal behaviours. In general, innocent suspects assumed different postures (uncrossed arms and legs, forward in the chair) and engaged in more frequent 'non-verbal' behaviours (movements of hands and postural changes) than guilty suspects; they also engaged in more smiles and head nods than guilty suspects. Innocent suspects were found to use more words per response, and were also more likely than guilty suspects to use words and phrases which were descriptive and reinforcing.

The findings in this study showed that in real-life decision-making both verbal and non-verbal behaviours observed during a BAI can be used in a complementary way. When these behaviours are observed in the context in which they occur they can be meaningfully assessed by persons who are trained and experienced in making these observations.

Additional research has indicated that training in some components of the BAI can enhance detection accuracy (Blair, 2007; Blair & McCamey, 2002). Both of these studies involved having participants view videotapes from the Horvath et al. (1994) study, and both found accuracies that exceeded 70 per cent after training was conducted.

Field study limitations

In all field deception or 'lie detection' research the ground-truth criterion problem is a difficult and complex methodological concern, and some have in large part dismissed the field research on the BAI because of this issue (Vrij et al., 2006). However, while the use of 'confessions' as

the criterion for ground truth can be problematic in some circumstances, corroborated confessions have been and continue to be the 'gold standard' in 'lie detection' field research, especially that dealing with polygraphy. Perhaps one of the major reasons for this is that it is difficult to establish reasonably certain ground-truth criteria in field applications without at least partial reliance on corroborated confessions (Honts, 1996; Patrick & Iacono, 1991). Nevertheless, research in which direct comparisons between confession-based and other ground-truth criteria has been done has not revealed any effect on outcomes. For example, Horvath (1977) compared the 'accuracy' of blind and trained evaluators on two groups of polygraphic data, one confession-verified and the other unverified; accuracy did not vary as a function of verification. Similarly, Krapohl, Shull, and Ryan (2002) found no difference in outcomes when they compared the accuracy of polygraph results on a sample of guilty persons who confessed to their crime with that on guilty persons who did not confess but whose guilt was otherwise established. In this study, by the way, the polygraphic data were objectively scored and uncontaminated by the criteria that established ground truth. Moreover, it is to be pointed out that even in those cases in which DNA analysis has been used to demonstrate a wrongful conviction, a confession by the actual perpetrator often plays a role in buttressing the forensic testing.

In addition, it is exceedingly unlikely that in the Horvath et al. (1994) study any of the confession-based suspects were not actually 'guilty' or 'innocent' (when their 'innocence' was confirmed by the 'guilty' person in the same investigation). There are a number of reasons for this. First, false confessions are rare in 'real-life' cases. While some experimental findings may suggest otherwise, the best evidence, based on three different approaches to estimation of the

rate of actual false confessions, shows that rate to be quite low. The first approach relies on national level data to estimate the number of wrongful convictions produced by false confessions annually. This requires an estimate of the number of convictions annually, the error rate in those convictions, and the proportion of errors due to false confessions. Using this method, two United States-based studies estimated the annual rate of false confessions ranged between 0.001 per cent to 0.04 per cent for all FBI index crime convictions (Cassel, 1998; Huff, Rattner, & Sagarin, 1986). Some observers have rejected this method as being too error prone to produce accurate estimates (Kassin, 1997; Leo & Ofshe, 1998).

The second method of estimating the frequency of false confessions involves identification of such confessions in a random sampling of criminal cases. Cassell (1998) did such a search in 173 cases that were in the sampling frame of a previous study; he did not find any false confessions in that sample (Cassell & Hayman, 1996). Other studies, designed to examine the interrogation process, did not mention the presence of any false confessions in their sample of cases (Irving, 1980; Leo & Ofshe, 1998; Softley, 1980). It would be assumed that the presence of a false confession would have been noted, especially so since the authors of one of these studies have also done research on false confessions.

A third method involves surveying samples of court-convicted criminals who are asked to self-report false confessions. Using this approach, Gudjonsson and Sigurdsson (1994) surveyed 95 per cent of all inmates entering prison in Iceland in a one year period ($n = 229$). They found that while 12 per cent claimed to have made a false confession in the past, none acknowledged a false confession to the current offence. In another study, Sigurdsson and Gudjonsson (1996) surveyed 509 prison

inmates and 108 juvenile offenders in Iceland. None of the juveniles claimed to have made a false confession. Approximately 12 per cent of the adult inmates said they made a false confession to the police in the past; less than 1 per cent claimed to have falsely confessed to the current offence. Gudjonsson and Sigurdsson (1994) also noted that about 92 per cent of all Icelandic prison inmates confess to the crime for which they are convicted. Based on such information, Gudjonsson (2003) estimated that the rate of false confessions in Iceland is below 1 per cent, and approximately half of these are interrogation produced.

Finally, collections of false confession cases may also provide some insight here. If false confessions are common, a large number of cases should be available for analysis. Yet, the most comprehensive survey of such cases in the United States found only 125 proven cases of false confession in the last 30 years (Drizin & Leo, 2004). Because it is so difficult to establish ground truth in criminal cases, this number is probably an underestimate of the actual number of false confessions. Nonetheless, if the estimate is off by a factor of 1,000, the rate of interrogation-induced false confessions in criminal cases in the US would still be less than 1 per cent. Moreover, it should be noted that some of the cases identified by Drizin and Leo as involving false confessions actually involved only admissions which might imply guilt, but were not confessions of guilt (Blair, 2005).

Each of the methods of estimating the frequency of false confessions has strengths and weaknesses. Taken together, however, the available data clearly show that false confessions are rare occurrences in actual criminal cases.

The second reason that it is unlikely that false confessions occurred in the Horvath et al. (1994) study is that all of the suspects were involved in private investigations, not police matters. The suspects were not in

custody and had the right to leave at any time; therefore, the interrogators in the Horvath et al. study had significantly less power over the suspects than is common in police interrogations. Third, the confessions that occurred (all to thefts) led to resolution of the incidents wherein the 'innocent' were exonerated and the 'guilty' were dealt with as their employers thought appropriate. A false confession would have been obvious and immediately called to the attention of the investigative firm as the confessions were corroborated and found to be consistent with the facts of the case. Fourth, the confessions in these cases did not result from the kind of 'high pressure' interrogational tactics that have been demonstrated to cause false confessions in actual criminal cases (Blair, 2005). It is confessions in those situations, not private investigations of organisational thefts, which have generally led to the concern about false confessions. Finally, as in similar research, the confessions establishing ground truth had been corroborated by other evidence or circumstances; they did not stand alone.

Additionally, if it is assumed that most of the confessions were true, then a single false confession (or even several false confessions) would have served only to lower the power to find differences between the behaviours of the guilty and innocent. In other words, in the unlikely event that the group of 'confessors' contained persons who falsely confessed, say, in order to protect a co-worker, it would be expected that the inclusion of such an innocent suspect (incorrectly identified as guilty) would reduce, not improve, the likelihood of finding statistically significant differences in behaviour between the guilty and the innocent.

Factual analysis is a somewhat less used form of ground-truth verification. For this reason the process is described here in some detail. Factual analysis refers to the expert

processing of the known information pertaining to an investigation (Inbau et al., 2001; Jayne & Buckley, 1999). In the Horvath, Jayne, and Buckley (1992) and Horvath et al. (1994) studies, factual analysis was carried out by two experts who independently assigned probabilities of 'guilt' or 'innocence' in each of five separate categories for each suspect; these were: biographical, opportunity/access, behaviour/attitudes, motivation/propensity and evidence. From the evaluators' separate evaluations a total probability of innocence or guilt was calculated for each suspect.

All 87 suspects' investigations were subjected to factual analysis. Of the 36 suspects whose ground truth was established by confession or other independent evidence, only one produced final factual analysis scores from both evaluators which were greater than 90 per cent and which were also inconsistent with ground truth. That is, when there was an agreement by both evaluators at levels of 90 per cent or higher, all but one of the confession/evidence verified suspects were correctly classified. Therefore, requiring at least a 90 per cent confidence level from both evaluators for including non-confession verified suspects appeared to provide a satisfactory criterion when ground truth could not otherwise be reasonably established. In this way, 24 suspects were included in the sample whose status was confirmed by factual analysis; 15 of these were 'innocent' and 9 were 'guilty'. In other words, the Horvath et al. (1994) sample of suspects included 36 persons whose 'guilt' or 'innocence' was confession-confirmed; 24 who were confirmed by factual analysis. Even though the criterion used to form the latter group is less certain as ground truth than is a confession, Horvath et al. (1994) reported: '... evaluators' scores on both the confession and the fact-analysis verified suspects were compared. This was done by separately subjecting each evaluator's scores for the various

assessments of suspects' behaviour to a two-way ANOVA ... These analyses did not reveal any significant effects for the Verification factor. Moreover, Chi-square tests showed no relationship between Verification and the frequency of correct, wrong or inconclusive judgments' (p. 801). In other words, the method of verification did not influence either the scoring of the data by the evaluators or the accuracy of the outcomes of their scorings.

Because false confessions are unlikely to occur in actual criminal cases and because corroboration of a confession with the facts of the case provides a safeguard against false confessions, confessions provide the strongest indicator of ground truth that is available in most field research. The use of factual analysis as an indicator of ground truth is clearly less certain and therefore less convincing. However, the similarity of the findings using the two types of ground-truth criteria in the Horvath et al. (1994) study suggests that factual analysis may be an effective substitute when confessions are not available. Additionally, multiple indicators of a construct are often used in research when, as was the case in the Horvath et al. study, they are internally reliable. This increases confidence in the outcome.

Experimental laboratory studies

Two experimental, laboratory-based studies involving issues directly or indirectly relating to the BAI have been reported. The first of these, with only an indirect relationship to some aspects of the BAI, was done by Kassin and Fong (1999). In the first phase of this study, participants either committed a mock crime or engaged in another similar, but non-criminal, activity. The guilty participants were given an incentive not to confess to the crime. All of the participants were then interviewed (while being video tape recorded) by a civilian who pretended to be a detective. This interview began with the suspect signing a waiver of his or her

Miranda rights (ie to legal representation, etc). Next, the interviewer explained to the participant how he or she had become a suspect, and asked the suspect to account for time when the crime was committed. Third, the interviewer refused to accept a suspect's denials, pounded his fist on the desk, and demanded the truth. Fourth, the interviewer attempted to break down the suspect's story, and finally, the interviewer asked the suspect to sign a confession. None of the suspects actually signed the confession. The interviews ranged from 3.5 to 6 minutes in duration.

During the second phase of the study, participants were randomly assigned to either a training or no-training group. Those in the training group watched two training videos developed and promoted by John E. Reid and Associates, Inc. These were video tapes designed to improve the ability to detect deception from observations of behavioural cues. Following the training, all of the participants, those in the trained and untrained groups, then assessed the tapes that were produced during the interviews with the suspects in phase one. Each person in each group judged whether or not the suspects in the videos were innocent or guilty. The reported results showed that the untrained evaluators were more accurate than the trained persons. Based on this finding the authors concluded that exposure to the Reid training tapes did not improve the accuracy of 'lie detection' and may have actually reduced it.

A second laboratory-based study, one in which the BAI was directly evaluated, was conducted by Vrij et al. (2006). In this study, the subjects were randomly divided into a guilty group, who committed a mock crime, or an innocent group, who did not commit a crime. The innocent subjects completed a complex set of tasks which involved playing a game, 'Connect Four'. While playing the game, they observed several people entering and leaving the room.

Later, a person entered the room looking for a lost wallet, and the subject helped in a search for the wallet. The 'new' person then left the room, and after a brief time returned to claim that money was missing from the wallet.

The deceptive subjects in this experiment simply entered the room where the wallet was located, removed money from the wallet, and then, as instructed, concocted a story about playing 'Connect Four' using a detailed written account of the scenario provided by the experimenters. Persons in both groups were motivated by an offer of money if they were assessed as truthful regarding the theft of money from the wallet. Importantly, all of the participants were recruited for this experiment by being told that it was about telling a convincing story.

Following completion of their instructed actions, the participants were interviewed by a uniformed police officer. The interview began with some introductory questions and then the suspect had a chance to deliver an open-ended narrative about the event that was 'experienced'. Following that, the interviewer concluded the interview by asking the suspect an array of 15 behaviour-provoking questions commonly included in a BAI (Horvath et al., 1994).

The 15 questions were then transcribed and assessed by coders. Each question was coded by two trained coders on two 5-point scales for the verbal assessments. The results pertaining to these assessments indicated that the answers given by the truthful participants to most of the BAI behaviour-provoking questions were more likely to be coded as deceptive than the answers given by deceptive participants. In other words, the behaviour-provoking questions in this study produced findings regarding message content opposite those said to occur in real-life BAI interviews.

Assessments of non-verbal behaviours in the Vrij et al. (2006) study were coded on

only three of the 15 BAI questions. Vrij et al. reported that these results indicated that the coding of these behaviours was generally in the opposite direction from what has been reported by Inbau et al. (2001).

Laboratory experiment limitations

While the primary limitation in field studies is the difficulty in establishing ground truth, the primary limitation of laboratory studies is external validity. That is, even though 'ground truth' is known with certainty in laboratory studies, the results of a given study may not accurately reflect what occurs in the real world. To limit this uncertainty it is necessary to design laboratory studies that closely mimic the key aspects of the real world phenomenon that is of interest. Both of the studies discussed above failed to do this.

The Kassin and Fong (1999) study, as already noted, did not directly assess the effectiveness of the BAI. It was instead an evaluation of the use of behavioural indicators in an 'interrogation', something which would only occur subsequent to, not be a part of, an actual BAI in the field. However, because the behaviours that were assessed in this study were said to be those advocated by Reid and Associates, it is appropriate to consider these findings in the context of the BAI. As was discussed in the general overview of the BAI section, Inbau et al. (2001) explicitly state that an interview must be non-judgmental and non-accusatory if the behaviours of the interviewee are to be assessed accurately. Since the Kassin and Fong study was intentionally designed to include an accusatory component, it demonstrates a clear lack of understanding of the Reid BAI process and as such it is clearly unjustified to use either behavioural models that were designed as a part of the BAI or to generalise from those results to what happens in an actual BAI.

The external validity of the Vrij et al. (2006) study is also questionable. While the

interviews were not accusatory, the scenario used by Vrij et al. did not resemble the typical real-life scenario in which the BAI is applied. First, the assumptions described earlier regarding the Sherlock Holmes Effect were not met. That is, the context of the Vrij et al. study was one in which all of the 'suspects' were involved in a situation in which they were isolated from other 'suspects'. This is very much unlike a real-life matter. In an actual case, say, for example, one in which money was stolen from a co-worker's wallet, all of the suspects are in the work environment together; they would know each other; would have developed relationships and opinions about each other; would know who has access to what areas; would know who might have the need to commit such a theft and so forth. When they are asked questions such as: 'Who do you suspect?'; 'Who do you think would have had the best opportunity to commit this theft?'; 'Who would you vouch for and eliminate from suspicion?'; they would have a context to use to frame their answers. In the Vrij et al. study this was not true. The design did not provide enough of a context for the participants to provide meaningful answers to the behaviour-provoking questions. Additionally, the participants were not given enough time to consider the situation (exactly how much time they were given is unclear as the Vrij et al. paper says only that the participants were given a few minutes). As was previously discussed, the suspect must have considerable time (at least a few hours) to consider what happened, who might have done it, and discuss this information with others if the Sherlock Holmes Effect is to be observed. It is also unclear whether or not a scenario in which the participants are aware that they are being assessed on their ability to 'tell a convincing story' can generate the concern on the part of the innocent persons necessary to produce the Sherlock Holmes Effect. This type of scenario merely creates

a situation in which deception is a game with a cash prize. This is certainly not the situation in actual interviews of criminal suspects. It should also be noted that in the truthful condition there was nothing for the suspects to be suspicious about as no money was actually stolen from the wallet.

It is not clear what type of realistic situation Vrij et al. (2006) attempted to mimic; having guilty suspects make up a story about playing 'Connect Four' with another participant rather than actually involving themselves in an event similar to that in which the innocent participants engaged is not a common real-life experience. A design that more accurately reflected actual investigations could have been done. For example, having both the innocent and guilty suspects play 'Connect Four' while the guilty suspect attempted to steal the wallet would be a possibility. The Vrij et al. study also confounded any interpretation of the results. That is, having the theft of the wallet co-vary with the fabrication of a story does not permit 'guilt' for the incident to be sorted from the 'lying' about the story. Whether the results were as they were because of the stealing or because of the fabrication is not possible to know.

The coding procedure utilised by Vrij et al. (2006) also represents a substantial misunderstanding of the BAI process. Vrij et al. incorrectly stated that non-verbal behaviour is evaluated on only three of the behaviour-provoking questions. In fact, non-verbal behaviour is evaluated on all of the behaviour-provoking and other questions that are asked during the BAI process (Inbau et al., 2001, pp. 126–127). Vrij et al. also separated the coding of non-verbal from the coding of the verbal process even though Inbau et al. clearly state that these behaviours must be considered in conjunction with each other (p. 126). Additionally, Vrij et al.'s representation of the way that non-verbal behaviours are evaluated during

the BAI is inaccurate. While specific behaviours (such as adaptors) are sometimes used as examples, non-verbal behaviours are considered in the context of the BAI assessment criteria discussed previously. A behaviour is not considered to be indicative of guilt or innocence until it has been assessed in the light of these criteria. In the Vrij et al. study all of the non-verbal behaviours that occurred during the three questions that were evaluated were coded in isolation from other criteria.

While the coders in the Vrij et al. (2006) study generally scored the verbal responses reliably, it is impossible to know whether their scorings would be consistent with what experienced BAI interviewers would assign. Information on the training of coders in Vrij et al. is sparse, and it is possible that the findings regarding the non-verbal behaviours would not generalise to a properly conducted BAI. It should also be noted that because only the behaviour-provoking questions were analysed, the study was really an evaluation of only the behaviour-provoking questions and not the BAI as it is commonly practised and described in the literature (Inbau et al., 2001).

CONCLUSION

There is a clear need for additional evaluations of the BAI, both field- and laboratory-based. It is hoped that this paper provides a grounding for that research to move forward in two ways. First, the Sherlock Holmes Effect, its assumptions, and its relationship to the BAI process have been explicitly described. The reason this has not been done before is elementary (the bad pun is not intentional). The Inbau et al. (2001) manual was intended to be a training tool, written by practitioners for practitioners. As such there was little concern with or need to consider the underlying 'theory' and the associated assumptions. In

reality the assumptions of the Sherlock Holmes Effect are almost always met in actual investigations; practitioners simply assume that this is understood. However, researchers do not necessarily make such assumptions, and the Sherlock Holmes Effect may not be given consideration in laboratory studies. It should, however, be seen as one of the critical design elements in any experiment that purports to test the BAI. Second, it is hoped that the overview of the BAI presented in this paper, and the detail given regarding the available field and laboratory assessments, will provide researchers with a better understanding of the BAI process and of the interrelationship of verbal and non-verbal indicators as they are synthesised by field practitioners. Directly dealing with these and related concerns will strengthen and, hopefully, encourage more laboratory-based studies in the future.

Well-designed laboratory studies can provide useful information. In the part of the 'lie detection' field dealing with instrumental methods of 'detecting deception', considerable attention has been devoted to the question of how best to simulate real-life conditions in the laboratory. It is generally recognised that in the latter environment it is difficult to replicate the motivation and the perceived consequences of the real world. In fact, it is precisely that difference between those two environments that has been the foundation for the controversy regarding how best to interpret the empirical evidence in polygraphy (Faigman, Kaye, Saks, & Sanders, 2002; Lykken, 1998). Nevertheless, there are conditions which, if maintained in the laboratory, do appear to optimise detection and to be advantageous to replicating the real-life environment (Kircher, Horowitz, & Raskin, 1988). In addition, research has shown that there may in fact be only differences in degree, not kind, between good

laboratory-based results and field data (Kircher, Raskin, Honts, & Horowitz, 1988; Pollina, Dollins, Senter, Krapohl, & Ryan, 2004).

Given these findings in polygraphy it appears likely that continued research on widely used methods applied in the real world to distinguish between 'innocent' and 'guilty persons', such as the BAI, will similarly reveal the optimal way to structure laboratory conditions to approximate those in the real world closely. This, of course, will enhance our understanding of the BAI and of the factors which influence its effectiveness.

It is especially important to produce convincing designs in this area that give specific attention to the interrelationship between gestures and speech, content and context, as is done in the real-life BAI. Such research should inform and perhaps alter the behaviour of practitioners. For this to occur, however, practitioners must be convinced that the research is relevant to their practice. Given the available laboratory studies on the BAI, there is no reason at this time for them to hold this conviction. Further research, properly carried out and reported, will be welcomed, especially by practitioners who, on a daily basis, have to address the common but difficult task of 'lie detection'.

REFERENCES

- Blair, J. P. (1999). *Detecting deception: The effects of the Reid Behavioural Analysis Interview training*. Unpublished master's thesis, Western Illinois University.
- Blair, J. P. (2005). A test of the unusual false confession perspective using cases of proven false confessions. *Criminal Law Bulletin*, 41, 127–144.
- Blair, J. P. (2007). Can detection of deception response bias be manipulated? *Journal of Crime and Justice*, 29, 141–152.

- Blair, J. P., & Kooi, B. (2004). The gap between training and research in the detection of deception. *International Journal of Police Science & Management*, 6, 77–83.
- Blair, J. P., & McCamey, W. P. (2002). Detection of deception: An analysis of the behavioural analysis interview technique. *Law Enforcement Executive Forum*, 2, 165–170.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10, 214–234.
- Cassell, P. G. (1998). Protecting the innocent from false confessions and lost confessions and from Miranda. *Journal of Criminal Law and Criminology*, 88, 497–556.
- Cassell, P. G., & Hayman, B. S. (1996). Police interrogation in the 1990s: An empirical study of the effects of Miranda. *UCLA Law Review*, 43, 839–931.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behaviour link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893–910.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129, 74–118.
- Drizin, S. A., & Leo, R. A. (2004). The problem of false confessions in the post-DNA world. *North Carolina Law Review*, 82, 891–1007.
- Ekman, P., O'Sullivan, M., & Frank, M. (1999). A few can catch a liar. *Psychological Science*, 10, 263–266.
- Ericson, R. (1981). *Making crime: A study of detective work*. Toronto: Butterworth and Company.
- Faigman, D., Kaye, D., Saks, M., & Sanders, J. (2002). *Modern scientific evidence: The law and science of expert testimony* (2nd ed.). Eagan, MN: West.
- Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. Hoboken, NJ: Wiley.
- Gudjonsson, G. H., & Sigurdsson, J. F. (1994). How frequently do false confessions occur? An empirical study among prison inmates. *Psychology, Crime and Law*, 1(1), 21–26.
- Honts, C. R. (1996). Criterion development and the validity of the Control Question Technique in field application. *Journal of General Psychology*, 123, 309–324.
- Horvath, F. S. (1973). Verbal and nonverbal cues to truth and deception during polygraph examinations. *Journal of Police Science and Administration*, 1, 138–152.
- Horvath, F. S. (1977). The effect of selected variables on the interpretation of polygraph records. *Journal of Applied Psychology*, 62, 127–136.
- Horvath, F. S., & Jayne, B. P. (1990). *A pilot study to quantify behavioural indicators of deception in structured interviews of criminal suspects* (Report prepared for US Department of Defense. Contract M503–89–07).
- Horvath, F. S., Jayne, B. P., & Buckley, J. P. (1992). *Trained evaluator's judgments of behavioural characteristics of truthful and deceptive criminal suspects during structured interviews* (Report of grant #904–90–C–1164, US Department of Defense).
- Horvath, F. S., Jayne, B. P., & Buckley, J. P. (1994). Differentiation of truthful and deceptive criminal suspects in behavioural analysis interviews. *Journal of Forensic Sciences*, 39, 793–807.
- Horvath, F. S., & Meesig, R. (1996). The criminal investigation process and the role of forensic evidence: A review of empirical findings. *Journal of Forensic Sciences*, 41, 963–969.
- Huff, C. R., Rattner, A., & Sagarin, E. (1986). Guilty until proven innocent: Wrongful conviction and public policy. *Crime and Delinquency*, 32, 518–544.
- Inbau, F. E., Reid, J. E., & Buckley, J. P. (1986). *Criminal interrogation and confessions* (3rd ed.). Baltimore, MD: Williams and Wilkins.
- Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogation and confessions* (4th ed.). Gaithersburg, MD: Aspen.
- Innes, M. (2003). *Investigating murder: Detective work and the police response to criminal*

- homicide*. Oxford, UK: Oxford University Press.
- Irving, B. (1980). *Police interrogation: A case study of current practice* (Royal Commission on Criminal Procedure). UK: HMSO.
- Jayne, B. P., & Buckley, J. P. (1999). *The investigator anthology*. Chicago, IL: John E. Reid and Associates.
- Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist*, *52*, 221–233.
- Kircher, J., Horowitz, S., & Raskin, D. (1988). Meta-analysis of mock crime studies of the control question polygraph technique. *Law and Human Behaviour*, *12*, 79–90.
- Kircher, J., Raskin, D., Honts, C., & Horowitz, S. (1988). Generalizability of mock crime laboratory studies of the control question polygraph technique. *Psychophysiology*, *25*, 462–463.
- Krapohl, D. J., Shull, K. W., & Ryan, A. (2002). Does the confession criterion in case selection inflate polygraph accuracy estimates? *Forensic Science Communications*, *4*(3), 1–10.
- Leo, R. A., & Ofshe, R. J. (1998). The consequences of false confessions: Deprivations of liberty and miscarriages of justice in the age of psychological interrogation. *Journal of Criminal Law and Criminology*, *88*, 429–496.
- Lykken, D. (1998). *A tremor in the blood: Uses and abuses of lie detection*. New York: Plenum Trade.
- Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology*, *89*, 137–149.
- Patrick, C. J., & Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, *76*, 229–238.
- Pollina, D. A., Dollins, A. B., Senter, S. M., Krapohl, D. J., & Ryan, A. H. (2004). Comparison of polygraph data obtained from individuals involved in mock crimes and actual criminal investigations. *Journal of Applied Psychology*, *89*, 1099–1105.
- Reid, J. E., & Arther, R. (1953). Behaviour symptoms of lie detector subjects. *Journal of Criminal Law, Criminology, and Police Science*, *44*, 106–112.
- Reid, J. E., & Inbau, F. E. (1977). *Truth and deception: The polygraph ("lie-detector") technique* (2nd ed.). Baltimore, MD: Williams and Wilkins.
- Sigurdsson, J., & Gudjonsson, G. (1996). Psychological characteristics of 'false confessors': A study among Icelandic prison inmates and juvenile offenders. *Personality and Individual Differences*, *20*, 321–329.
- Simon, D. (1991). *Homicide: A year on the killing streets*. Boston, MA: Houghton Mifflin.
- Softley, P. (1980). *Police interrogation: An observational study in four police stations* (Royal Commission on Criminal Procedure). UK: HMSO.
- Vrij, A., Mann, S., & Fisher, R. P. (2006). An empirical test of the behavioural analysis interview. *Law and Human Behaviour*, *30*, 329–345.