Wiley Series in the Psychology of Crime, Policing and Law

Detecting Lies and Deceit

Pitfalls and Opportunities

Second Edition

WILEY

Aldert Vrij

Detecting Lies and Deceit

Second Edition

Wiley Series in

The Psychology of Crime, Policing and Law

Series Editors

Graham Davies and **Ray Bull**

University of Leicester, UK

The Wiley Series in the Psychology of Crime, Policing and Law publishes concise and integrative reviews on important emerging areas of contemporary research. The purpose of the series is not merely to present research findings in a clear and readable form, but also to bring out their implications for both practice and policy. In this way, it is hoped the series will not only be useful to psychologists but also to all those concerned with crime detection and prevention, policing, and the judicial process. Current titles of interest in the series include:

Offender Profiling: Theory, Research and Practice
Edited by Janet L. Jackson and Debra A. Bekerian

Psychology, Law and Eyewitness Testimony
Peter B. Ainsworth

Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice
Aldert Vrij

Children's Testimony: A Handbook of Psychological Research and Forensic Practice
Edited by Helen L. Westcott, Graham M. Davies and Ray H.C. Bull

Stalking and Psychosexual Obsession: Psychological Perspectives for Prevention, Policing and Treatment
Edited by Julian Boon and Lorraine Sheridan

The Psychology of Interrogations and Confessions: A Handbook
Gisli H. Gudjonsson

Terrorists, Victims and Society: Psychological Perspectives on Terrorism and its Consequences
Edited by Andrew Silke

Reducing Crime: The Effectiveness of Criminal Justice Intervention
Edited by Amanda Perry, Cynthia McDougall and Dave P. Farrington

Practical Psychology for Forensic Investigations and Prosecutions
Edited by Mark Kebbell and Graham Davies

Detecting Lies and Deceit

Pitfalls and Opportunities

Second Edition

Aldert Vrij

John Wiley & Sons, Ltd

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, ONT, L5R 4J3, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available is electronic books.

Library of Congress Cataloging-in-Publication Data

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-470-51624-9 (ppc) 978-0-470-51625-6 (pbk)

Typeset in 10/12pt Century Schoolbook by Aptara Inc., New Delhi, India
Printed and bound in Great Britain by TJ International, Padstow, Cornwall
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

About the Author

Aldert Vrij is a Professor of Applied Social Psychology in the Department of Psychology at the University of Portsmouth (UK). He has published more than 300 articles and book chapters to date, mainly on the subjects of nonverbal and verbal cues to deception and lie detection. He also advises the police about conducting interviews with suspects, acts as an Expert Witness in court, and gives invited talks and workshops on lie detection to practitioners and scholars across the world. He is at present Editor of *Legal and Criminological Psychology*, a forensic journal published by the British Psychological Society.

Series Preface

The Wiley Series on the Psychology of Crime, Policing and the Law publishes reviews of important areas of contemporary research. The purpose of the series is not only to present research findings in a clear and readable form, but also to bring out their implications for both policy and practice. In this way, it is hoped that the series will not only be of use to psychologists, but also to all those concerned with crime detection and prevention, policing and judicial processes.

The current volume has a focus on detecting lies and deceit and therefore also on detecting truth. The opening chapter importantly deals with widely held myths concerning detecting deception and the (consequent) errors people make when trying to decide if someone is being deceptive or truthful. The next chapter notes that many people often lie (if only for the best of intentions). Thus, most people may learn how to lie successfully. The third and fourth chapters present knowledgeable, in-depth overviews: (a) of the various psychological theories that have been offered concerning nonverbal and verbal behaviours/cues which may relate to deception/truthfulness and (b) of the cues found by research actually to be related to deception. The fifth chapter reviews the growing body of research on people's beliefs about which cues they think are indicative of lying and the origins of the many beliefs that the research presented in the previous two chapters reveals to be false beliefs. A comprehensive chapter follows on how good people are at detecting truth/lies, including relevant professionals.

Chapters 7 to 10 examine a variety of procedures/techniques that have been developed to try to determine deception/truthfulness from people's behaviour and speech. A number of these procedures, including some widely marketed around the world, have not been found by published research to be effective. Indeed, reliance on them could result in the innocent being persecuted and the guilty remaining at large to carry out further wrong-doing.

Chapters 11 to 13 focus on attempts to detect deception from bodily/brain activity. Here, the available research again confirms that many such procedures achieve levels of performance not far from, or close to, chance. These chapters and chapter 14 also seek to explain why catching liars is often so difficult.

The crucial closing chapter focuses on how people might improve their lie detection skills, not in the ill-informed way that many books on catching liars claim, but in an in-depth way informed by a wealth of psychological theory and research.

In this broad and deep second edition (that cites close to 1100 references) Professor Aldert Vrij remarkably has managed to produce a book of even greater quality than his (2000) first edition (which had an impact world-wide). It is an example of scholarship at the highest level. This book will help people realise why it is usually so difficult to discriminate lying from truth-telling – honestly it will!

RAY BULL
and
GRAHAM DAVIES
August 2007

Preface

When I published my *Detecting Lies and Deceit* book in 2000, I did not anticipate publishing a second edition for at least a decade. However, important events have changed my mind. Since 2000, the world has experienced several terror attacks and security threats, and the United States started its "War on terror". Against this background, the desire to detect lies has become more urgent than ever before and governments are calling upon scholars to design lie detection tools to protect their citizens from attack.

Scientists have responded by conducting research, and at present more than 150 articles on deception and lie detection are published in peer-reviewed journals each year. Several scholars have claimed that they have developed lie detection tools that are highly accurate, and they encourage governments to use their tools. What do they propose? Is it substance, is it spin, or perhaps a mixture of both? My view is that substance and spin are often intertwined. I also believe that the only way to distinguish between the two is by providing a comprehensive overview of facts regarding lying and lie detection. A factual account enables the reader to generate an informed opinion about how accurate these claims are.

This book provides a comprehensive review of deception research published to date. It is different from the previous edition in several aspects. First, the old text is updated and hundreds of new studies have been added. Second, it also contains descriptions of lie detection methods that were not discussed in the 2000 edition, such as the Behaviour Analysis Interview, Scientific Content Analysis (SCAN), Voice Stress Analysis, thermal imaging, P300 brain wave research, and functional Magnetic Resonance Imaging (fMRI) research. Third, because so much has changed in the world of deception research, it was essentially easier to mostly rewrite my book than try to add in new developments. I therefore rewrote and restructured most of the old text. For example, this book provides overview chapters regarding why professionals often fail to detect deceit and what they can do to become better lie detectors. Such chapters were not included in the 2000 edition.

The result, I believe, is a book about deception that is better structured and more comprehensive than the 2000 edition. I also realise that this book contains considerably more text than the 2000 edition, and so perhaps readers will prefer to dip into individual chapters rather than read the book from beginning to end. To facilitate this, I have ensured that the chapters can be read independently from each other.

Acknowledgements

I am very grateful to Dr Samantha Mann for her valued comments on previous drafts of this book. I also would like to thank Dr Sharon Leal for her insightful comments on the physiological chapters (Chapters 11 to 13) and Chapter 15, and Professor Pär Anders Granhag for his constructive help with Chapter 15. I also wish to express gratitude to the institutions that have funded my research in the past and at present. These are in alphabetical order: the British Academy, the Dutch Government, the Dutch Organisation for Scientific Research (NWO), the Economic and Social Research Council (ESRC), the Leverhulme Trust, the Nuffield Foundation, the UK Government, the US Government and the University of Portsmouth. Without their financial support, I could not have carried out my research and would not have written this book. In many of my studies the participants were police officers from the Netherlands and the United Kingdom. I highly appreciate their willingness to participate, because it gives invaluable insight into the views and skills of people who detect lies on a professional basis. I am also grateful to the police for giving us access to videotapes of police–suspect interviews. This gave us the opportunity to conduct unique research into the behaviour of truth tellers and liars in realistic high-stakes situations. Finally, I am grateful to John Wiley & Sons for publishing a second edition of this book.

CHAPTER 1

Introduction

MYTHS IN DECEPTION

Everybody knows what it is to lie, and this familiarity with lying surely makes us knowledgeable about deception.[1] For example, we all know that lying is undesirable. We therefore rarely lie ourselves since decent people do not make a habit of lying, and this lack of practice makes us poor liars. Because lying is so undesirable, we prefer not to be lied to. We therefore do not wish to spend time with liars and try to avoid them.

Fortunately, most people we know do not make good liars either. They reveal their deceit by behaving nervously and avoiding eye contact. Hence, just by observing someone's behaviour we can often spot a liar. We are rather good lie detectors, particularly when spotting mendacity in our own children, partners, and close friends.

Of course, conmen, smugglers, and other types of criminals try to achieve their aims by means of deceit, and, if successful, they could do us a lot of harm. Fortunately, we are well protected against them, because professional lie catchers are good at spotting such liars. The additional benefit these professionals have is that they can rely upon specialised lie detection equipment. Indeed, some of the machines that were used in the past may have proven to be unreliable, but a lot has changed

[1] I will use the terms lying and deception interchangeably.

since then. For example, technological developments such as brain scanners mean that researchers have now direct access to people's thoughts and feelings and therefore can tell with certainty whether someone is lying. Moreover, professionals nowadays use much more sophisticated interrogation techniques to catch liars than they did in the past.

Correct? Actually. . . . No. All of the above statements mentioned here are myths rather than facts, and I will unravel those myths in this book.

GOOD LIARS AND POOR LIE DETECTORS

People tend to *under*estimate their own *ability to lie* (Elaad, 2003). There are several reasons why people think they are worse liars than they actually are. First, they tend to overestimate the extent to which their own thoughts, emotions, and other mental states are transparent to others (e.g., *illusion of transparency*, Gilovich, Savitsky, & Medvec, 1998). In other words, people mistakenly believe that their lies shine through. Second, self-perceptions are typically characterised by positive illusions (Taylor & Brown, 1988), and people typically think of themselves as being more moral than most others (Kaplar & Gordon, 2004). Admitting to being a good liar does not match with this positive self-image. Third, although people tell more white lies than serious lies, they remember their serious lies better than their white lies (Elaad, 2003). This book shows that it is often somewhat easier to detect serious lies than white lies, so people mostly remember the lies that probably were more easily detected (serious lies). Finally, perhaps people remember better those times when their lies failed than when they lied successfully, particularly if these failed lies resulted in negative consequences. By forgetting the times they lied successfully, they may underestimate how often their lies succeed.

People tend to *over*estimate their own *ability to detect lies* (Elaad, 2003) and more lies remain unnoticed than they generally think. There are many reasons why lies remain undetected, and they can be clustered into three main categories (Vrij, 2007): poor motivation; difficulties associated with lie detection; and common errors made by lie detectors.

Poor Motivation: The Ostrich Effect

One reason why lies remain unnoticed is that often people do not attempt to detect them, because they do not want to learn the truth. I label this phenomenon the *ostrich effect*. There are at least three reasons why someone might not want to know the truth. First, a fabrication might sometimes sound more pleasant than the truth, and in such cases ignorance might be preferable. For example, why bother trying to discover whether mendacious compliments about one's body shape, hairstyle, dress sense, or achievements and so on, are truthful?

Second, people sometimes do not investigate whether they have been lied to because they fear the consequences the truth may hold. Some statistics suggest that up to 50% of men and 40% of women engage in extramarital relations (Feldman & Cauffman, 1999), yet many of them will remain undiscovered. A husband, for example, might try to dismiss suspicions that his wife is having an affair and avoid discovering the truth, because if he were to discover the truth and confront her with his knowledge about her lover, she may decide to leave her husband. This may be something he does not want her to do. Therefore, communicating what he has discovered may have undesirable consequences for the betrayed man and, upon realising this, he may decide not to investigate the issue. After the scandal with Monica Lewinsky broke, President Clinton told his aides in the White House that he did not have a sexual relationship with her. Erskine Bowles, the White House Chief of Staff at that time, was more than willing to believe him. This is how he described that moment to the Grand Jury: "All I can tell you is: This guy who I've worked for looked me in the eye and said he did not have sexual relationships with her. And if I didn't believe him, I couldn't stay. So I believe him" (*The Independent*, 14 September 1998, p. 4).

Third, people sometimes do not want to detect lies because they would not know what to do if they came to know the truth. Most guests, for instance, will not try to find out whether their host is truthful in his claims that he likes their presents, because what would they do if they discovered that he did not like their presents? More serious lies may remain undetected for the same reason. Suppose that the husband's wife in the example above decides not to leave him, what should he himself do instead? Once betrayed, the cuckolded husband may have trouble trusting his wife again, and the repercussions of such a discovery may take a very long time to resolve. President Clinton's personal secretary, Betty Currie, tried to avoid learning details of the relationship between President Clinton and Monica Lewinsky probably for the same reason. On one occasion, Lewinsky said to Currie of herself and the President: "As long as no one saw us – and no one did – then nothing happened." Ms Currie responded: "Don't want to hear it. Don't say any more. I don't want to hear any more" (*The Observer*, 13 September 1998, p. 8). Indeed, it is clear to see that knowledge of the affair would put Ms Currie in a difficult situation, forcing her between publicly declaring this knowledge or acting as an ally, and that is why she probably preferred to remain ignorant.

Difficulties Associated with Lie Detection

As I will demonstrate in this book, even when people try to detect lies, they often fail to do so. Research has indicated that even professional lie catchers, such as customs officers and police officers, often make incorrect decisions, and that their ability to separate truths from lies typically does not exceed that of laypersons. One category of reasons why even motivated people fail to catch liars is because lie detection is difficult. Perhaps the main difficulty is that, as this book reveals, not a single nonverbal, verbal, or physiological response is uniquely associated with deception. In other words, the equivalent of Pinocchio's growing nose does not exist. This means that there is no single response that the lie detector can truly rely upon. Another difficulty is that liars who are motivated to avoid being caught may attempt to exhibit nonverbal, verbal, or physiological responses that they believe make an honest impression on lie detectors. As we will see, liars who employ such so-called *countermeasures* can indeed often fool professional lie detectors.

Common Errors Made by Lie Detectors

Another category of reasons why people fail in their attempts to catch liars is that they make errors. In this book I will discuss numerous errors that lie detectors commonly make, including a tendency to pay attention to cues that are not reliably associated with deception. One reason as to why they focus on those non-diagnostic cues is because they are taught to do so. For example, Inbau and his colleagues wrote an influential handbook about how to interrogate suspects. The most recent, fourth edition, of this book appeared in 2001 (Inbau, Reid, Buckley, & Jayne, 2001).[2] In their book, they provide information about how lying suspects usually behave. According to Inbau and colleagues, behavioural cues to deception include posture changes, gaze aversion, self-adaptors (stroking back of head, touching nose, straightening or stroking hair, pulling threads on clothing and so on), placing a hand over the mouth or eyes when speaking, and hiding hands (by sitting on them) or hiding feet (by pulling them under the chair). In particular the beliefs that liars place their hands over the mouth or eyes or that they avert their gaze are frequently mentioned in the police literature (Brougham, 1992; Gordon & Fleisher, 2002; Kuhlman, 1980; Macdonald & Michaud, 1992; McKinnon, 1982; Rabon, 1992; Walkley, 1985; Walters, 1996; Waltman, 1983; Yeschke, 1997; Zulawski & Wicklander, 1993). As I will

[2] Although there are many police manuals, I will mainly focus on the Inbau, Reid, Buckley, and Jayne (2001) manual, because their manual is commonly used by police and military interrogators and hence is so influential (Gudjonsson, 2003).

demonstrate, these behavioural cues to deception, and most of the other behavioural cues to deception mentioned by Inbau *et al.*, are not identified as such in the existing deception literature. There is evidence that lie detectors who pay attention to these cues actually perform worse than those who do not (Kassin & Fong, 1999; Mann, Vrij, & Bull, 2004).

LACK OF REALISM

People's tendency to be overconfident in their ability to detect lies becomes evident when reading the deception literature or listening to practitioners. In principle, lies can be detected via observing someone's behaviour, analysing their speech, or measuring their physiological responses. In all three areas practitioners and researchers can be found who make bold claims about their ability to detect lies that they fail to back up with research findings. Several examples illustrate this. Paul Ekman, an American emeritus professor of psychology, has specialised in nonverbal cues to deceit. His work has inspired academics and practitioners for several decades. Not long ago, he claimed that his system of lie detection can be taught to anyone with an accuracy of more than 95% (*New York Times Magazine*, 5 February 2006; see also *Washington Post*, 29 October 2006 for a similar statement). There is, however, no study published to support this claim (Chapter 6).

One of the interview techniques discussed in detail in Inbau *et al.*'s (2001) manual is the Behaviour Analysis Interview (BAI). This manual is linked to a training programme where John E. Reid and Associates teach practitioners the BAI method and other methods. On their website, http://www.reid.com/training_programs/r_interview.html, they report that more than 300,000 professionals in the law enforcement and security fields have attended their three-day programme since it was first offered in 1974. They further claim that interviewers specifically trained and experienced in behaviour analysis assessment can correctly identify the truthfulness of a person 85% of the time (Inbau *et al.*, 2001). However, conclusive evidence to support this claim is lacking (Chapter 7, see also Blair & Kooi, 2004).

Udo Undeutsch, a German emeritus professor of psychology, laid the foundations for a verbal lie detection tool nowadays called Statement Validity Assessment (SVA). He reported that the method has been applied in thousands of criminal cases investigating child sexual abuse in Germany and Sweden and that in no single case has the outcome been later contradicted by other relevant evidence (Undeutsch, 1982). This suggests that the method is highly accurate. However, such a premise is not supported by SVA research. In the SVA chapter I discuss an alternative reason for why the verdicts of SVA experts are unlikely to be contradicted: there is often no factual evidence available in child sexual abuse cases, and therefore often no opportunity to falsify the claims made by SVA experts (Chapter 8).

The debate about the use of the polygraph as a lie detector is heated, as will be explained in this book. Faith in the accuracy of the polygraph is high among practitioners. Dan Sosnowski, an American polygraph examiner, claimed that evaluation of the polygraph in the US has showed that it detects deception with 97% accuracy (*The Independent*, 11 October, 1999, p. 1). Sosnowski's claim is exaggerated and strong faith in the accuracy of the polygraph can be questioned on the basis of the scientific polygraph literature (Chapter 11).

Pavlidis, Eberhardt, and Levine (2002a) reported in the prestigious journal *Nature* that they have developed a high-definition thermal-imaging technique that can detect attempted deceit by recording thermal patterns around the eyes. They suggested that the technique has the potential to be used in rapid security screening without the need for skilled staff or physical contact. Unsurprisingly, their article attracted considerable media attention, because it sounds promising: it implies that the device could be used at airports to detect smugglers and potential terrorists. However, in their subsequent erratum Pavlidis, Eberhardt, and Levine (2002b) fell somewhat short from this claim by mentioning that "it was not intended to convey the impression that this thermal imaging technique is *already* suitable for mass security-screening purposes" (italics added by me). As I will explain in this book, it is doubtful whether lies could ever be reliably detected in the way suggested in the *Nature* article (Chapter 11).

Not long after the introduction of measuring brain responses, claims have been made that these techniques can be highly successful in catching liars. For example, Dr Farwell developed a "brain fingerprinting" technique that he aggressively promotes on his company's website www.brainwavescience.com. He claims high accuracy in classifying truth tellers and liars by measuring their brain waves. Others are less convinced. Wolpe, Foster, and Langleben (2003) reported that relatively few participants were actually tested and most of the data regarding brain fingerprinting are not published in peer reviewed literature.[3] The

[3] Peer review is a crucial part of scientific writing. Peer review is a process whereby submitted articles are scrutinised by the editor and at least two consultants. Those consultants are scientists typically working in the same field as the author(s), and they are therefore knowledgeable about the topic of investigation and the research methods described in the paper. Most articles submitted to peer review get rejected because the editor and consultants detect some serious flaws. Those who do not get rejected typically undergo thorough revision before they are published.

latter point means that rigorous scientific scrutiny of Farwell's technique has not yet taken place. To back up his claim of high accuracy, Farwell refers on his website to his study with Smith published in 2001. In their well-documented report, the National Research Council (2003, p. 162) discuss this study and note that "the range of stimuli to which examinees were exposed were very small, and the sample size was very small ... Whether these findings generalise to other, more complex contexts in larger groups is not known." Interesting is the view of Emanuel Donchin, a psychophysiologist working at the University of South Florida, and Farwell's former graduate adviser and co-author on one of Farwell's articles. Reflecting on Farwell's technique he reports that "The necessary research has never been done" (Knight, 2004). In other words, the accuracy of brain fingerprinting is not as well established as Dr Farwell suggests (Chapter 12).

Ruben Gur examines activity in brain structures and areas. He is part of a research group that claims that their test is 99% accurate in distinguishing truths from lies and ready to detect terrorists (Wild, 2005). The 99% accuracy study has not been published in a peer reviewed journal, and the accuracy rates that have been published in such journals to date are lower. Moreover, the premise behind the lie detection technique Gur promotes could be too simplistic. Gur claims that "a lie is always more complicated than the truth" (Wild, 2005), which is not necessarily the case, as this book reveals (Chapter 13).

THE CONTENT OF THIS BOOK

This book discusses nonverbal, verbal, and physiological indicators of deceit and the ability to detect lies on the basis of these indicators. In particular, I will address two questions: (1) Are there systematic differences between truth tellers and liars in nonverbal behaviour, speech content, and physiological responses? And (2) to what extent are observers able to distinguish between truths and lies when they examine someone's nonverbal behaviour, speech content, and physiological responses?

I commence this book with general information about deception. After defining deception, I describe which types of lie exist, the reasons why people lie, estimates of how often people lie, and individual differences in telling lies. Chapter 2 shows that lying is very much part of everyday life, and that the role of lying in social communication is two-pronged: sometimes lying causes harm to the ones who are lied to, but many lies told in daily life are white lies that may even benefit the lie receivers, often serving as a social lubricant.

Chapter 3 describes the relationship between nonverbal behaviour and deception. I present several theories as to why nonverbal cues to deception may emerge, and discuss research examining the behaviours displayed by truth tellers and liars. This includes research about the behaviours exhibited by suspects in police interviews and by the politicians Bill Clinton and Saddam Hussein. The chapter reveals that most nonverbal cues appear not to be associated with deception and some nonverbal cues are, at best, only weakly related to deception. It further shows that the relationship between nonverbal behaviour and deception is complicated, amongst other reasons because different people show different cues to deception, and because a liar's behaviour depends on the context in which the lie takes place.

Chapter 4 is the first of five chapters examining verbal cues to deception. The verbal cues discussed in later chapters (Chapters 7 to 10) form part of existing verbal lie detection tools used by professional lie catchers and scholars. In Chapter 4 I summarise verbal cues that are not part of such existing tools. This chapter shows that some of these verbal cues show weak relationships with deception.

In Chapter 5 I discuss how people believe liars behave and what they think they say. This chapter reveals that people often hold erroneous views about nonverbal and verbal cues to deception. Many cues that people believe are indicators of deceit are actually not related to deception, whereas some cues that they do not associate with deceit are in fact weakly related to deception. This chapter also shows that these erroneous views are held across the world and by both laypersons and professional lie catchers, such as police officers, customs officers, immigration officers, and prison officers.

Chapter 6 discusses how accurate laypersons are in detecting truths and lies in people they do not know, or in their friends, romantic partners and children, when they rely on the alleged liar's nonverbal and verbal behaviour. The chapter shows that people are typically poor at this task under such circumstances, even if it concerns the truths and lies told by friends and relatives. This chapter further discusses how good professional lie detectors are in distinguishing truths from lies when they pay attention to someone's nonverbal and verbal behaviour. The chapter shows that they typically do not fare better than laypersons.

Chapters 7 to 13 discuss the various lie detection tools used by professionals and scholars. Chapter 7 deals with the Behaviour Analysis Interview (BAI), the only professional lie detection tool to date that examines nonverbal cues to deception. BAI, however, also examines some verbal cues. BAI is taught to practitioners all over the world. In Chapter 8 I describe Statement Validity Assessment (SVA), a verbal veracity assessment tool developed in Germany and Sweden to assess the credibility of statements made by alleged victims of sexual abuse. To date, SVA is the most widely used verbal veracity detection instrument and SVA assessments are accepted as evidence in criminal courts in several countries. Chapter 9 introduces Reality Monitoring (RM), another verbal veracity assessment tool. As far as I know, RM is not used by practitioners but it is popular amongst scholars, perhaps because it has a strong theoretical foundation. In Chapter 10 I discuss Scientific Content Analysis (SCAN). Like BAI, SCAN is taught worldwide.

Chapters 11 to 13 deal with physiological cues to deception. Throughout history it has been assumed that lying is accompanied by physiological activity within the liar's body. This activity is nowadays measured in different ways, mostly with a machine called a polygraph, also referred to as a lie detector (this labelling, however, is misleading as I discuss in Chapter 11). For lie detection purposes the polygraph measures finger sweating, blood pressure, and respiration. There are different theoretical rationales as to why truth tellers and liars may show different physiological responses while being attached to the polygraph. These different rationales lead to different interview protocols. Concern-based interview protocols are discussed in Chapter 11 and orienting reflex-based interview protocols in Chapter 12. Although concern-based interview protocols are used worldwide, the use of orienting reflex-based interview protocols is restricted to mainly Japan and Israel.

In Chapter 11 I also discuss thermal imaging, a technique that measures the blood flow around the eyes, and voice stress analysis, a technique that measures tremors and other aspects of the voice, both of which are concern-based lie detection techniques; and in Chapter 12 I also discuss a technique based on the orienting reflex that measures P300 brain waves via electroencephalograms (EEG-P300). Such alternatives techniques to polygraph testing are sometimes presented as fundamentally different from polygraph testing, but such claims are misleading. They measure concern or the orienting reflex differently from the polygraph, but use the same concern-based and orienting reflex-based interview protocols as employed in polygraph testing. Therefore, these techniques share the strengths and weaknesses associated with these two types of interview protocols.

In Chapter 13 I describe lie detection based on measuring activity in brain structures and areas. These activities are measured with a functional magnetic resonance imaging (fMRI) brain scanner. This is the most recent development in physiological lie detection, but, again, not fundamentally different from polygraph testing. The fMRI lie detection tool employs the same concern-based and orienting reflex-based interview protocols as used in traditional polygraph testing.

A review of the scientific literature discussed in Chapters 7 to 13 reveals that truth and lies are detected at levels well below perfection with each of these veracity assessment tools. However, it also shows that with several of these techniques truths and lies can be detected more accurately than by simply observing someone's nonverbal and verbal behaviour.[4]

Chapters 3 to 13 show several pitfalls in lie detection. I present these pitfalls in a systematic manner in Chapter 14 and provide 15 reasons as to why people fail to catch liars. The final chapter (Chapter 15) deals with improving lie detection skills. I make some suggestions about how to improve lie detection via analysing speech or measuring physiological responses. I will argue that collaboration between deception researchers and researchers in other areas of psychology is therefore needed. Most of Chapter 15 deals with how to improve lie detection via observing someone's nonverbal or verbal behaviour, and I present 17 guidelines that lie detectors could use to detect deceit in this manner. In this chapter I pay more attention to this type of lie detection than to that based on speech analysis or physiological responses and I do this for two reasons. First, it is the most inaccurate type of lie detection, and therefore perhaps mostly in need of improvement. Second, this form of lie detection can be used in many more situations than the other two types of lie detection, because it does not require transcribing someone's speech (necessary for many speech analysis protocols) or equipment such as a polygraph, a cap containing electrodes (often used to record EEGs), or fMRI scan.

[4] When I refer to verbal lie detection via analysing speech in a systematic manner using the verbal veracity assessment tools discussed in Chapters 7 to 10, I will use the term *speech analysis*. When I refer to observing verbal cues in an unsystematic manner, I use the term observing (or paying attention to) *verbal behaviour*.

CHAPTER 2

Lying: A Selfish Act and a Social Lubricant

Deceiving others is part of everyday life. A man says that he is pleased with his birthday presents, although in fact they are not what he really wanted; the host receives compliments about his cooking, although the food was not really good; and a schoolgirl watching TV tells her dad that she has finished her homework, although she has not actually yet started it. Moreover, a boy insists that he hasn't taken money from his mother's purse; a smuggler tells the customs officer that she has nothing to declare; and a murderer denies having committed the crime. These are only a few examples from an almost inexhaustible source of possible lies.

This chapter deals with telling lies in everyday life. After defining deception, I describe the different types of lie that exist, the reasons why people lie, estimates of how often people lie, and individual differences in telling lies. This chapter demonstrates that on average we each try to dupe others more than once a day. It further shows that the role of lying in social communication is two-pronged: Sometimes it is harmful for the ones who are lied to. The world may look entirely different today, if people had been fully aware of Hitler's real intentions before World War II broke out; if people realised that those individuals who flew airplanes into New York's Twin Towers took flying lessons in order to fulfil just that purpose; or if people realised that those who carried out the bombings in London on 7 July 2005, in effect, led double lives prior to these attacks. However, many other lies we tell are white lies, and are harmless to those who are duped (e.g., "Of course will you soon find a new boyfriend", "I think you are much prettier than her", "I think you performed really well"). As this chapter reveals, these kinds of lies may even benefit the lie receivers ("targets") and often serve as a social lubricant. I will demonstrate that lying is an important phenomenon in interpersonal relationships and we often like the company of people who lie frequently. This chapter therefore challenges the stereotypical view that "lying is always bad".

A DEFINITION OF DECEPTION

Deception can be defined in many ways. Mitchell's (1986, p. 3) definition in particular is remarkable. He defines deception as "*a false communication that tends to benefit the communicator*." This is a broad definition and classifies many acts as deceit. For example, from this perspective some plants can lie. Bond and Robinson (1988) describe how the orchid *Ophrys speculum* dupes male wasps with the illusion of sexual contact. The orchid produces an odour that mimics insect sexual pheromones to attract and arouse male wasps. Males are drawn to the centre of the flower, because the colouring resembles a female wasp. In the centre of the flower the male finds thick, long hairs that resemble the hair found on a female wasp's abdomen. The male wasp believes he has found a mate, and so a pseudo-copulation takes place. The wasp then moves on to another orchid, and the orchids are cross-pollinated in the process.

Mitchell's definition is controversial because it implies that unconsciously and mistakenly misleading others should also be classified as deception. Suppose a boy told his mother that he gave her all the change from the shopping but unwittingly left behind a 50p coin in his pocket. The boy is lying according to Mitchell's definition. Similarly, a sales assistant who has not been informed by her boss that a product's price has been reduced, and who therefore asks for too much money, is lying according to Mitchell's definition. However, many people will disagree with this. Peterson (1995) examined whether people tend to classify mistakes as lies. Participants (both five year olds and adults) were given the following story. A boy witnessed an event involving a green car and was then asked later by his father whether he still remembers the colour of the car. The boy, who thought he remembered but really had forgotten the colour of the car, replies: "Yes Daddy, I remember. The car was yellow". Participants were asked whether they thought the boy in the story was lying. The vast majority, 88% of five year olds and 95%

of adults, thought the boy was not lying. For many people a core issue of deception is that it is an *intentional* act. Many people will thus not think that the orchid in the example above is lying while duping the male wasps.

Following Krauss (1981), many researchers therefore define deception as "*an act that is intended to foster in another person a belief or understanding which the deceiver considers to be false*" (Zuckerman, DePaulo, & Rosenthal, 1981, p. 3). A boy who misremembers the colour of a car and therefore states the wrong colour is not lying, whereas a boy who actually remembers the colour of the car but deliberately states the wrong colour, is lying. The delusional man who believes that he is Napoleon is not lying in his claims. It is obvious that he is not telling the truth, but he believes his own story and so does not have any intention to deceive other people. Similarly, a woman who mistakenly believes that she was sexually abused in her childhood and who therefore goes to the police to report this abuse, has given a false report, but is not lying. It is crucial to distinguish such so-called false beliefs from lying, as it can be very difficult to detect false beliefs while paying attention to behaviour, speech content, or physiological responses (Ceci & Bruck, 1998). I return to this issue in more detail in later chapters in this book.

Box 2.1 Misremembering emotional events

There is growing evidence that people are able to "remember" emotional incidents that never occurred. Although very young children may be disproportionately vulnerable to this kind of memory error (Ceci, Huffman, Smith, & Loftus, 1994; Ceci, Loftus, Leichtman, & Bruck, 1994), adults make such errors too (Porter, Yuille, & Lehman, 1999). In the Porter *et al.* (1999) study, 77 undergraduate students were interviewed. During these interviews, they were presented with various events (e.g., falling on their head, getting a painful wound, or being sent to a hospital emergency room). They were told that, according to their parents, these events had occurred in their childhood. The interviewer gave further details about the events supposedly given by the parents. Unknown to the interviewees, the events were invented by the researchers and had never happened to the participants according to their parents. Guided imagery instructions were given to the participants to help them generate images for the false event (e.g., "Visualise what it might have been like and the memory will probably come back to you"). Results indicated that 26% of students "recovered" a complete memory for the false event, and another 30% recalled aspects of the false experience.[1]

[1] Research suggests that a distinction needs to be made between memories recovered during therapy versus spontaneously recovered memories, see Geraerts (2006) and Geraerts *et al.* (2006, 2007).

According to Krauss' definition, sarcastic remarks are not lies either. Indeed, someone who makes a sarcastic remark deliberately does not tell the truth, but this form of lie-telling is not meant to foster a false belief in another person. On the contrary, the deceiver wants the deception to be uncovered, and will try to make this clear to another with either facial expressions or by changing their tone of voice (Zuckerman, DeFrank, Hall, Larrance, & Rosenthal, 1979).

Defining lying as an intentional act also implies that if two people contradict each other, this does not automatically mean that one of them is lying. In criminal cases, different people who have witnessed the same event may recall the event somewhat differently from each other, and sometimes their statements may even be contradicting. It could thus be that none of the witnesses are lying but that at least one of them misremembers the event.

Lying does not necessarily require the use of words. The athlete who fakes a foot injury after a bad performance and limps off the track is lying without using words. It is also possible to lie by hiding information, although, again, this must happen intentionally. Taxpayers who deliberately do not report a particular source of income on their tax return are lying, whereas those who forget to give this information are not lying. The man who tells his wife who was present at the dinner party is not lying when he accidentally forgets to mention his secretary. However, he is lying when he consciously does not mention her.

Another aspect of deception (apart from being an intentional act) is that it is defined solely from the perspective of the deceiver and not from the factuality of the statement. A statement is a lie if the deceiver believes what he or she says is untrue, regardless of whether the statement is in fact true or false. An untruthful statement is therefore not necessarily a lie (e.g., the boy mentioned above who misremembers the colour of a car is not lying), and, also, an actual truth could be a lie. Suppose that a suspect, who believes that his friend is hiding in his apartment, tells the police that his friend is abroad. This statement is a lie even when, unknown to the suspect, his friend has actually fled the country. Ironically, those truthful statements could be associated with "cues to deceit", as I will explain in Chapter 3.

Defining lying from the perspective of the deceiver further implies that a statement that was initially a lie could lose that status over time (Pickel, 2004; Polace, 2004; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001). For example, in Pickel's (2004) study, participants were instructed to tell the truth or lie about details of a film they had watched. A week later they were asked to accurately recall what they had seen. The participants who were asked to lie the week before gave more incorrect details than the participants who recalled the details truthfully the first time. Also, some of the participants reported as accurate details that they had concocted for the interview a week before, suggesting that they now had come to believe their own fabrications. For those witnesses their untruthful statements were no longer lies.

Burgoon and Buller (1994, pp. 155–156) defined deception slightly differently from Krauss. They define deception as "*a deliberate act perpetrated by a sender to engender in a receiver beliefs contrary to what the sender believes is true to put the receiver at a disadvantage*". The main difference between this definition and Krauss' definition is in the last seven words "*to put the receiver at a disadvantage*". This extension of the definition is unfortunate. There are times when people tell lies to make targets appear better or to protect their feelings rather than to put targets in a disadvantageous position. This will be discussed later in this chapter.

However, Krauss' definition is not entirely satisfactorily either, because it ignores another aspect of deception. Ekman (1985/2001) argued that people are only lying when they do not inform others in advance about their intentions to lie. Magicians are therefore not lying during their performance, as people in the audience expect to be deceived. Also, liars sometimes fail to mislead their target, despite having clear intent to do so. For example, the target may know that the information the liar wants him or her to believe is untrue. In these cases, the attempt to deceive the target has failed, but such unsuccessful attempts can still be classified as lies. I therefore prefer to define deception as "*a successful or unsuccessful deliberate attempt, without forewarning, to create in another a belief which the communicator considers to be untrue*."

People sometimes fool themselves – a process called self-deception. After failing an examination, students often delude themselves into believing that they were not motivated enough to revise thoroughly for the exam, rather than acknowledging that they do not really understand the topic. Self-deception has both negative and positive features (Lewis, 1993). On the one hand, people can ignore or deny the seriousness of several bodily symptoms, such as a lump in a breast or a severe chest pain during physical exertion. This can be life-threatening. On the other hand, self-deception may serve to protect self-esteem. After being rejected on a date, people may convince themselves that they were not interested in dating this person after all, instead of admitting that their affection was not reciprocated. According to my definition, deception is an act that involves at least two people. This definition excludes self-deception, which will therefore not be further discussed in this book.

According to my definition, nonhumans can also lie. DeWaal (1986, also described in Bond & Robinson, 1988) gives some examples of sophisticated animal deceit, including a case about chimpanzees' bluffing. Male chimps use mutual bluff displays to find out which is the strongest. Sometimes chimps show involuntary teeth-baring during these displays, which is a sign of fear. Obviously, teeth-baring, if perceived, would undermine a chimp's bluff. Chimps therefore tend to turn their back before baring their teeth and resume their bluffing once the expression has gone. In one instance, DeWaal even observed that a chimp quickly used his fingers to push his lips back over his teeth. This book, however, solely deals with human deceit.

TYPES OF LIES

Psychologists distinguish between outright lies, exaggerations, and subtle lies (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996).[1] Outright lies (also referred to as falsifications, Ekman, 1997) are lies in which the information conveyed is completely different from what the liar believes to be the truth. A guilty suspect who denies any involvement in the crime, is telling an outright lie; so is a student who says that she was revising for her exam when, in fact, she was shopping. Moreover, applicants who claim in a selection interview that they are happy in their present job but just fancy a change after having been in that job for so many years, are telling an outright lie when the real reason for their application is that they have been sacked; and when he testified in the Paula Jones[2] case, former US President Clinton told an outright lie when he said that he could not remember whether he had ever been alone together with Monica Lewinsky in the Oval Office.

Exaggerations are lies in which the facts are over- or understated (understatements are also called minimisations; Tyler, Feldman, & Reichert, 2006). People can exaggerate their regret at arriving too late

[1] Non-psychologists use similar categories but label them differently. For example, Metts (1989) calls outright lies, exaggerations, and subtle lies "contradictions", "distortions", and "omissions", respectively.

[2] Paula Jones was a former government employee who accused Clinton of demanding oral sex from her in his hotel room in Little Rock, Arkansas, in 1991.

at an appointment with a friend; can embellish their remorse for committing a crime during a police interview; can present themselves as more diligent than they in fact are during a job interview; or can underplay the expense of their shopping purchases in a conversation with their partner.

Subtle lying involves literal truths that are designed to mislead. Clinton was telling such a lie in 1999 when he said to the American people that he "did not have sexual relations with that woman, Miss Lewinsky." The lie was subtle, because the statement implied that nothing of a sexual nature had happened between the two of them, whereas he was relying on the narrower definition that they did not actually have sexual intercourse. Another type of subtle lying involves concealing information by evading the question or omitting relevant details. Someone who does not like a picture painted by her friend can conceal the extent of her opinion by saying that she does like the bright colours used in the painting. Passengers who describe to customs officers the contents of their luggage are concealing information if they have illegal drugs in their possession that they deliberately fail to mention.

When romantic partners lie to each other they do so relatively often by concealing information (Metts, 1989). There are several reasons as to why liars prefer concealments. First of all, they are difficult to detect. Once information is provided, lie detectors can verify the accuracy of this information by searching for further evidence that supports or contradicts it. In the case of concealments, however, no information is given. Moreover, concealing information is relatively easy. When telling an outright lie or when exaggerating, a liar should invent a story that sounds plausible, whereas nothing needs to be invented when concealing information. Another problem with telling an outright lie or exaggerating is that liars need to remember the details they provided in case the topic of the lie comes up on subsequent occasions. However, they don't need to remember anything if they don't provide information (concealment). Also, if called upon to explain concealed information the liar may claim to have forgotten to mention it, or that (s)he didn't realise that (s)he hadn't mentioned it, but an excuse may be more difficult to concoct if an outright lie or exaggeration is discovered. Also, concealments are often perceived as less negative than outright lies (Bavelas, Black, Chovil, & Mullett, 1990; Levine, 2001; Levine, Asada, & Lindsey, 2003; Spranca, Minsk, & Baron, 1991), therefore the liar may face fewer repercussions from the lie target in case a concealment is detected than in case an outright lie is detected. Finally, concealments are probably easier to morally justify. Liars could justify concealments by thinking "Everything I said was the truth. . .I just didn't mention the *whole* truth".

An alternative typology of lies can be made by examining lie complexity and the consequences of getting caught in a lie (Ekman, 1985/2001; Ekman & Frank, 1993; Vrij, 1998). Some lies are more difficult to tell than others. For example, lying is more difficult when the liar is aware that the target has some evidence. Suppose a 12-year-old boy who smokes despite being forbidden to do so by his parents. When asked whether he smokes, it is more difficult for him to deny when his mother shows him the empty cigarette packet she found in the pocket of his trousers than when his parents don't have any evidence to prove his smoking habits. Lying is also more difficult when the target is suspicious. The adulterous wife will probably have more difficulties in hiding her clandestine affair when her husband is suspicious than when he is not. Finally, a lie is more difficult to tell when the liar has no opportunity to prepare the lie. When a woman is asked out by a particular man that she does not want to go out with it is more difficult for her to lie when his request takes her by surprise and she has to lie spontaneously than when she has anticipated his request for a date and hence prepared an excuse.

The consequences of getting caught in a lie vary. They are more serious for a murderer when during a police interview he denies having committed the crime, than for a girl who exaggerates the number of CDs she owns in a conversation with friends. And the consequences are bigger for smugglers when customs officers discover them with heroin than when they are in possession of too much alcohol. A classification on the basis of lie complexity and consequences is useful because liars may behave differently in different situations, as I will discuss in Chapter 3.

REASONS WHY PEOPLE LIE

When people think about lying, they typically think about lies that are told to gain material advantage or to avoid materialistic loss/punishment. These kinds of lies may be seen as selfish, disruptive of social life, and hurtful to the targets. This stereotypical negative view of lying is often expressed both in the popular press (DePaulo, Kashy *et al.*, 1996) and by ordinary people (Backbier, Hoogstraten, & Meerum Terwogt-Kouwenhoven, 1997; Bok, 1978; DePaulo, 2004; Kowalski, Walker, Wilkinson, Queen, & Sharpe, 2003; Robinson, 1994; Schweitzer, Hershey, & Bradlow, 2006).

However, people also lie for psychological reasons, and those lies typically receive more understanding (Seiter, Bruschke, & Bai, 2002). Goffman (1959) pointed out that life is like a theatre and that people often behave as actors on a stage. Following Goffman, many others have argued that the "self" that is presented to others in daily life is not the true self but an edited version (Allen & Gilbert, 2000; DePaulo, 2004; DePaulo, Kashy *et al.*, 1996, DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003; Feldman, Forrest, & Happ, 2002; McCornack, 1997; Turner, Edgley, & Olmstead, 1975; van Dongen, 2002). The edited self takes into account how one wishes to be seen by others. Thus people lie, for example, to avoid embarrassment: Who wants to reveal all their inadequacies, errors, and indecent and immoral thoughts to the rest of the world? People also lie to protect others: Who wants to tell the bold truth, thereby deliberately hurting the feelings of a good friend?

People do not automatically recognise the two-pronged nature of lying. Indeed, when participants in Backbier *et al.*'s study (1997) were asked about their attitude towards lying, they initially reacted by saying that "lying is bad", but after thinking in more depth about their attitude towards deception, they could report many instances in which they lied themselves, and showed understanding of their own lies. In a similar vein, although most participants in Boon and McCleod's (2001) study strongly endorsed the belief that complete honesty is important in a romantic relationship, many could envisage that in certain situations lying is justifiable.

Bella DePaulo and her colleagues conducted naturalistic deception research that demonstrated this two-pronged nature of lying (DePaulo & Bell, 1996; DePaulo & Kashy, 1998; DePaulo, Kashy, *et al.*, 1996; Kashy & DePaulo, 1996). They asked 77 American college students and 70 community members to keep a diary for seven days and to record all their social interactions and all of the lies they told during those interactions. A social interaction was defined as an exchange with another person that lasted 10 minutes or more. Brief encounters such as greetings in the morning were thus ignored. The diary study revealed that people lie for several reasons. These motives can be defined according to three dimensions (Vrij, 2007): (1) for one's own benefit (self-oriented) or for the benefit of others (other-oriented); (2) to gain advantage or to avoid costs; and (3) for materialistic reasons or for psychological reasons.

Examples of self-oriented lies to gain material advantage include a house owner who does not point out to a potential buyer all the shortcomings of the property that she is aware of, to attract a good offer for her house; and an applicant who exaggerates his current income during a selection interview to secure a higher income in his next job.

Self-oriented lies to gain psychological advantage are told to make a positive impression on others. For example, students may tell their parents that they just cleaned their room or could tell their friends that they knew most of the answers in a popular TV quiz.

We also tell self-oriented lies to avoid material loss/punishment. Those lies vary from guilty suspects who deny involvement in a crime during their police interviews, to young children who do not mention that they have secretly taken a biscuit.

Self-oriented lies to avoid psychological costs include lies to protect oneself from embarrassment. Children may be reluctant to admit that they are bullied at school because they feel ashamed or embarrassed at being singled out. When Bill Clinton admitted for the first time on television to the American people that he had been involved in an "inappropriate relationship" with Monica Lewinsky, the first reason he gave for having misled people was "a desire to protect myself from the embarrassment of my own conduct". Our research indicated that undergraduate students keep secrets to avoid psychological costs (Vrij, Nunkoosing, Paterson, Oosterwegel, & Soukara, 2002). They prefer to keep secret that they are still a virgin or that they are homosexual. They also sometimes deny to friends fancying a person they, in fact, do fancy.

The categories are not mutually exclusive and lies could fit more than one category. Sex offenders typically minimise the nature and extent of their sexual offending, not only because they want to avoid being sentenced, but also because they feel ashamed about their activities (Ahlmeyer, Heil, McKee, & Englis, 2000). Also, in the example above, Clinton may have lied about his relationship with Monica Lewinsky to avoid political recrimination.

We also tell other-oriented lies, which are lies told for other people's benefit (materialistic gain), to make others appear better (psychological gain), to avoid another person's materialistic loss/punishment, or to save the other person from psychological damage. Examples are a man telling a potential buyer of his friend's car, that the car is truly reliable when he knows that it is not (materialistic gain); a woman making her friend exaggerated or false flattering comments ("You are witty and attractive, of course you will soon find a new boyfriend!") to boost the friend's self-confidence (psychological gain); an innocent father owning up to a crime that he knows his son committed to save his son from a conviction (to avoid punishment of his son); or a woman pretending to have experienced an orgasm (to avoid upsetting her boyfriend).

People frequently tell what I call *social lies*. Conversations could become awkward and unnecessarily rude, and social interactions, including friendships and romantic relationships, could easily turn sour if people were to tell each other the truth all the time. In order to maintain a good working relationship with colleagues it is better to pretend to be busy when invited for lunch than to admit that you find their company boring and would rather avoid them. Similarly, a husband would do better to refrain from explicitly criticising his wife's dress when he knows that she bought the dress in a sale and hence cannot return it to the shop; and it may be kinder to respond with enthusiasm when receiving an expensive present from a friend even when you don't like the gift. Social relationships benefit from people making each other compliments now and again because people like to be liked and like to receive compliments (Aron, Dutton, Aron, & Iverson, 1989; Curtis & Miller, 1986). In that respect, social lies such as making deceptive but flattering comments ("I like your new haircut") may benefit mutual relations. Social lies are told for psychological reasons and serve both self-interest and the interest of others. They serve self-interest because liars may gain satisfaction when they notice that their lies please other people, or because they realise that by telling such lies they avoid an awkward situation or discussion. They serve the interest of others because hearing the truth all the time ("The steak you cooked was really tough"; "You look much older now than you did a few years ago"; "I would be surprised if you succeed in what you want to achieve", etc.) could damage a person's confidence and self-esteem.

FREQUENCY OF LYING

How often people lie is a difficult question to investigate. Probably the most popular method is to ask people to report how many lies they have told during a certain period of time (Cole, 2001; Arnett Jensen, Arnett, Feldman, & Cauffman, 2004; Vrij, Floyd, & Ennis, 2003). The question is how reliable are the outcomes? For example, are people really aware of the number of lies they tell? And are they willing to admit all their lies they remember, including the serious lies? Another method is to ask participants to write down the last lie they told (Backbier & Sieswerda, 1997). Again, there is no guarantee that people will actually remember the last lie they told, particularly if this was a trivial lie. Also, they may not be willing to share their last lie in case it was a serious lie. Yet another method is by showing participants hypothetical scenarios and asking them to indicate whether they would lie is such a scenario (Argo, White, & Dahl, 2006; Williams, 2001). With this method, there is no certainty that participants would actually respond in reality as they indicated they would in the scenario. Alternatively, people are asked to keep a diary of all the social interactions they have and to note how often they lied during these social interactions (DePaulo, Kashy *et al.*'s, 1996, diary study introduced above). Again, people may not be willing to report some of the serious lies they told. Also, instructing participants to report their social interactions may affect the nature of these interactions and may influence the number of lies they tell during these interactions. In a final method, participants are asked to hold a 10-minute get-acquainted conversation with another person. Unbeknown to the participants at the time of the conversation, the researchers measure the number of lies the participants told during these conversations[3] (Tyler *et al.*, 2006). A limitation of this method is that it only examines a 10-minute snapshot in an artificial laboratory setting.

Despite the problems related to the various methods, the findings of all studies measuring the frequency of deception converge to the same conclusion: lying is a frequent event.[4] For example, Tyler *et al.* (2006) found that in a 10-minute get-acquainted conversation session, participants lied on average 2.18 times. Moreover, 78% of the participants reported that they had lied during the 10-minute session. They lied, amongst others issues, about their academic achievements, finances, and personal skills. These lies fit in well with Goffman's edited-self perspective (e.g., "we often put on a show") introduced in the previous section. A survey reported in the British Sunday newspaper *The Observer* revealed that 88% of those who were interviewed have feigned delight on receiving a bad Christmas present and that 73% have told "flattering lies" about a partner's sexual ability (*The Observer*, OM Magazine, 11 January 2004, p. 12). These lies also fit well in the edited-self perspective.

The results of the diary studies (DePaulo, Kashy *et al.*, 1996) showed that lying is a fact of everyday life. Almost all participants admitted that they had lied during the week that they kept their diary. Undergraduate students reported telling two lies a day and community members told one lie a day. On average, participants lied in one out of every four of their social interactions, and they lied to 34% of all the people they interacted with over the week (remember, only interactions that lasted at least 10 minutes were recorded). More lies (around 50%) were self-serving than told in the interest of others (around 25%). Most lies (67%) were outright lies, and more lies were told for psychological reasons (around 60%) than for materialistic reasons (around 40%). In alignment with the edited self-perspective, participants lied often about their feelings, preferences, attitudes and opinions, and less often about their actions, plans and whereabouts. Lies about achievements

[3] The researchers measure deceit by asking participants to view the videotaped interview in private and identify any untruthful statements that they told.

[4] Bleske & Shackelford, 2001; DePaulo, Kashy *et al.*, 1996; Engels, Finkenauer, & van Kooten, 2006; Kalbfleisch, 2001; Knox, Zusman, McGinty, & Gescheidler, 2001; Lawson & Leck, 2006; Lippard, 1988; Mazur & Ebesu Hubbard, 2004; Turner *et al.* 1975; Tyler *et al.*, 2006; Weiss & Feldman, 2006; Williams, 2001; Whitty, 2002.

and failures were also commonplace (DePaulo, Kashy *et al.*, 1996). Participants reported that the overwhelming majority of the lies they told were not serious.[5]

Box 2.2 The perspective of targets

When I mention that liars do not perceive their lies as serious, I describe the perspective from the liar, and not from the perspective of the target of the lies. Research has addressed the question of how targets feel and react when they have discovered a lie told by their romantic partner. It appears that in the case of serious lies, lie tellers and targets react differently, with lie tellers interpreting lies in a more positive way than targets (Feldman & Cauffman, 1999; Kaplar & Gordon, 2004; Kowalski *et al.*, 2003). For example, lie tellers find their lies more altruistic and more justified than targets, whereas targets are more uncertain and confused about why they were lied to, or become more angry or sad. However, in an ingenious experiment, where the same scenarios were rated from different perspectives (the lie teller, target, and observer perspectives), it was found that both lie tellers and targets are biased and thus contribute to the difference between the two perspectives (Gordon & Miller, 2000). Compared to the observers, the lie tellers minimised the importance of the event, whereas the targets exaggerated the damage done by the lie.

There are also gender differences in targets' perspectives. Women are typically more distressed than men when they discover the deceit of a close relational partner (Levine, McCornack, & Avery, 1992) or find lying less acceptable and justifiable (Blair, Nelson, & Coleman, 2001; DePaulo, Ansfield, Kirkendol, & Boden, 2004; DePaulo, Epstein, & Wyer, 1993; Levine *et al.*, 1992). Men and women's reactions depend on the topic of the lie, however (Haselton, Buss, Oubaid, & Angleitner, 2005). Using a biological strategic interference perspective, it was hypothesised and found that women are more upset than men when members of the opposite sex lie about their economic resources, status, and depth of feelings towards the target, whereas men are more upset than women when members of the opposite sex lie about being available for sexual intercourse.

[5] When in another study, DePaulo and her colleagues examined the nature of the serious lies people tell, a different picture emerged (DePaulo, Ansfield, Kirkendol & Boden, 2004). Most of these lies were self-serving (90%) and were told to cover illegal and immoral acts.

Lawson (2000) investigated tellers' and targets' views not only about lies but also about truths. He found no differences between tellers and targets in evaluating little lies (e.g., "I like your drawing"), but tellers rated the truths ("I think this is a boring party") as less acceptable and more selfish than targets.

Discussing the frequency of lying is not straightforward, however, because it depends on several factors, such as to whom the lie is told, the situation, and personal characteristics of the liar. I will elaborate on these issues in the next three sections.

WHO DO PEOPLE LIE TO?

Further analyses of DePaulo's diary study, reported by DePaulo and Kashy (1998), revealed a relationship between telling lies and the emotional closeness of the relationship. By comparing the lies told between strangers, acquaintances, best friends and spouses, it was found that the highest rate of lying occurred in conversations with strangers, and the lowest rate occurred in conversations with spouses. One reason why people lie less to their romantic partners (and also to friends) than to strangers is that they have the desire to be honest to people they feel close to. However, there are also other reasons (Anderson, Ansfield, & DePaulo, 1999). The fact that friends and partners know more about us limits the topics that are suitable to lie about. We can try to impress strangers at a cocktail party by exaggerating our culinary skills but this is fruitless with friends who have experienced our cooking; and we can tell an acquaintance that we failed the exam because we did not spend enough time revising, but this will not work with our partners who saw us sweating over our books for many successive evenings. Also, people typically believe that those close to them can detect their lies. They therefore may believe that they will not to get away with the lie, and this may be another reason not to lie much in the first place.

Although people tend to lie less to those with whom they feel close, the frequency of lying between spouses was still far from incidental and occurred in nearly one out of every 10 social interactions (DePaulo & Kashy, 1998). Cole's (2001) study also showed evidence for the frequent nature of lying in romantic relationships. He reported that 92% of participants admitted lying to their significant other.

Some of the lies told between spouses are other-oriented lies and are told because the liars care about their partners and don't want to harm their feelings (Bell & DePaulo, 1996). Other lies told between spouses are self-oriented lies but are minor (DePaulo & Kashy, 1998). Perhaps a limited amount of trivial lying serves important privacy needs for individuals in close relationships (DePaulo & Kashy, 1998). However, interactions with spouses are also the domain of self-oriented serious lies. When people were asked to describe the most serious lies they had ever told to someone else, they overwhelmingly reported that the target of these lies were close relationship partners (Anderson *et al.*, 1999; DePaulo, Ansfield *et al.*, 2004). These lies were often told to cover serious issues, such as infidelities, and were told to save the relationship. Sometimes spouses believe that the truth cannot be told without threatening the relationship. In such instances, they may decide that lying is preferable. They perhaps do so reluctantly. They often feel uncomfortable while lying to their spouses (DePaulo & Kashy, 1998), but it is, in their view, the best option they have, given the circumstances (Boon & McLeod, 2001). Liars may well be correct in this assumption. Timothy Levine and his colleagues asked students to describe a recently recovered lie told by a romantic partner and their reactions upon discovery (Jang, Smith, & Levine, 2002; McCornack & Levine, 1990a,b). They found that almost 25% of the participants terminated their relationship after discovery of the lie. However, it was not the fact of lying that ended the relationship, but the information that the liar had tried to withhold from the target.

There are a few exceptions to the finding that people tend to lie less to those with whom they feel close. For example, undergraduate students often lie to their mothers (Backbier & Sieswerda, 1997; DePaulo & Kashy, 1998; Lippard, 1988). DePaulo and Kashy (1998) found that students lied in almost half of the conversations they had with their mothers. Perhaps they are still dependent on their mothers (e.g., monetarily) and sometimes have to lie to avoid disapproval and to secure financial resources (e.g., Darling, Cumsille, Caldwell, & Dowdy, 2006; Lippard, 1988). Another explanation is that they still care about what their mothers think of them. They therefore tell their mothers that they do not drink much beer, attend all lectures, study hard, and keep their room clean. A third reason for lying to mothers is that students want to achieve more autonomy and thus more independence from their parents (Arnett Jensen *et al.*, 2004; Mazur & Ebesu Hubbard, 2004).

SITUATIONAL FACTORS: A JOB AND A DATE

The regularity with which people lie also depends on the situation. Robinson, Shepherd, and Heywood (1998) interviewed undergraduate students, of whom 83% said they would lie in order to get a job. These students said that it was wrong to lie to best friends, but they saw nothing wrong in lying if it secured a job. They also thought that employers expected candidates to exaggerate qualities when applying. A MORI poll shed light on the scale of lying in job applications. Twenty-five percent of Britain's working population had misled their potential employers when applying for a job. Those lies ranged from providing false details of personal skills and qualities to exaggerated experience and salary (King, 2006; see also Birkeland, Manson, Kisamore, Brannick, & Smith, 2006; Levashina & Campion, 2006; Pontari & Schenkler, 2006; and Weiss & Feldman, 2006).

Rowatt, Cunningham, and Druen (1998) found that 90% of participants admitted being willing to tell a lie to a prospective date. About 40% of men and women indicated that they actually had told a lie to initiate a date with an attractive member of the opposite sex (Rowatt, Cunningham, & Druen, 1999). Also DePaulo's diary study revealed that people lied relatively often to their romantic partners in the early stages of their relationship (once in every three social interactions). One possible explanation is that people worry whether their "true self" is lovable enough to attract and keep these partners, and they therefore present themselves as they wish they are instead of how they actually are (e.g., the self-edited perspective, DePaulo & Kashy, 1998).

PERSONAL CHARACTERISTICS OF THE LIAR

Gender

Although no gender differences have been found in the frequency of lying (DePaulo, Kashy *et al.*, 1996), men and women tell different types of lies. Men tell more self-oriented lies than women, whereas women tell more other-oriented lies than men (DePaulo & Bell, 1996; DePaulo, Kashy *et al.*, 1996; Feldman *et al.*, 2002). In DePaulo and Bell's (1996) experiment, participants were asked to choose the two paintings they liked the most and the two paintings they liked the least from a selection of 19 paintings. After making their choice they wrote down what in particular they liked and disliked about each of the four paintings they had nominated. The participants were then informed that they would be meeting an art student with whom they would discuss their opinions about the four paintings. They were told that the art student may have actually painted some of the paintings herself, and she would tell the participant if she had. The participants were also told that the art student did not know which paintings the participant had included in their list.

In reality the art student was a confederate (a helper of the experimenter) who did know which paintings the participants liked and disliked. For each of the four paintings, the art student asked the participant what they thought of the painting overall, what they specifically liked about it, and what they disliked about it. As she opened the discussion about one of the paintings the participant liked, and about one that the participant disliked, the art student mentioned that the painting was one that she had painted herself. This created four different situations for each participant. In one discussion, they had to give their opinion about a painting they liked and which was not painted by the art student. In a second discussion they discussed a painting that they did not admire and which was not painted by the art student. In a third discussion they had to talk about a painting they liked and which was painted by the art student, and in a fourth discussion they discussed a painting by the art student that they did not like. Obviously, the most awkward discussion was the latter discussion where participants were forced to discuss the art student's painting that they did not like. The findings revealed gender differences in this situation in particular. The women exposed more positive feelings about the latter painting than the men did, especially by exaggerating their liking for the aspects of that painting that they really did like. Gender differences also emerged when participants were discussing the art student's painting that they really did like. Women exaggerated their liking of the painting to the artist more than men did.

DePaulo, Kashy *et al.* (1996) found in their diary study that women tell more other-oriented lies than men, particularly towards other women. In female–female conversations equal percentages of lies (around 40%) were other- and self-oriented, whereas in male–male conversations about six times as many self-oriented lies (around 60%) than other-oriented lies (around 10%) were told. Saarni (1984) has demonstrated that the female tendency to tell more other-oriented lies becomes evident in young children. Children aged seven to eleven received presents that they loved (such as sweets) as a reward for helping an adult with her work. Shortly after this, they were asked to help again, but this time the present they were given was boring and more suitable for younger children. Girls showed their disappointment to a lesser extent than boys and responded more enthusiastically when they received their dull present.

It has been shown that both men and women enjoy conversations with women more than conversations with men (DePaulo, Epstein, & Wyer, 1993; Reis, Senchak, & Solomon, 1985). Reis *et al.* (1985) provide several explanations for this phenomenon that are unrelated to deception. However, the finding that women tell more other-oriented lies than men, and thus make more flattering comments and more frequently avoid saying things which may hurt the other person, may also be a reason for why conversations with women are often appreciated.

Men are more willing than women to use deception in order to get a date (Eyre, Read, & Milstein, 1997; Rowatt *et al.*, 1998). Also, differences emerge in the types of lie men and women tell during a date (Benz, Anderson, & Miller, 2005; Eyre *et al.*, 1997; Keenan, Gallup, Goulet, & Kulkarni, 1997; Tooke & Camire, 1991; Whitty, 2002). Men tend to feign more frequently than women their earning potential (e.g., misleading members of the opposite sex about their career expectations), whereas women more frequently engage in deceptive acts to improve their physical appearance (e.g., "sucking in" their stomach when around members of the other sex). These deceptive acts reflect gender differences in preferred characteristics of potential partners. When 50 male and 50 female participants were asked what they look for in a potential partner, men were more likely than women to emphasise the importance of their partner's physical appearance, whereas women were more likely than men to emphasise the importance of their partner's earning capacity (Buss & Barnes, 1986).

Age

There is a debate in the literature regarding the age at which children start telling lies. There is agreement that children are capable of telling deliberate lies by the age of four (Bussey, 1992; Leekam, 1992; Newton, Reddy, & Bull, 2000; Wilson, Smith, & Ross, 2003). However, even before the age of four, children attempt to misinform others. In their experiment, Lewis, Stanger, and Sullivan (1989) instructed children aged between 33 and 37 months not to peek at a toy while the experimenter left the room. Most children transgressed and did look at the toy. When asked, the great majority of children either denied having peeked or would not answer the question. Nigro and Snow (1992), who replicated this study with children as young as 32 months, obtained similar results. Ceci and DeSimone Leichtman (1992) found that three year olds, each of whom believed that the interviewer was unaware that the child's mother broke a toy, frequently covered for her by claiming that it was broken by another or that they did not know who broke it. Chandler, Fritz, and Hala (1989), utilising a hide-and-seek board game, have shown that even two year olds withheld information by "erasing footprints", "laying false trails", and pointing to the wrong place in order to prevent another person from finding a treasure. Naturalistic studies revealed similar results. Newton *et al.* (2000) found that a two-and-a-half-year-old boy provided incorrect information in an effort to put his mother on the wrong track, and Reddy (2007) reported examples of deception that suggest that infants and toddlers are able to communicate false information as early as they are able to communicate true information.[6]

Several researchers suggest that the earliest lies children tell are those that are meant to escape punishment (Bussey, 1992; Lewis, 1993; Saarni, 1979; Stouthamer-Loeber, 1986; Wilson *et al.*, 2003). Lies generated to obtain rewards probably appear later (DePaulo & Jordan, 1982), followed by lies to protect one's self-esteem (Bussey, 1992). These are all self-oriented lies. There is evidence that even young children tell other-oriented lies. For example, Ceci and DeSimone Leichtman (1992) found that children as young as three years old will lie to protect a loved one, and Lewis (1993) gives an example of a three-year-old girl who responded with great enthusiasm when receiving a present from her grandmother, although she in fact did not like the present.

Children also lie when they are asked to do so. Tate, Warren, and Hess (1993) conducted an experiment in which an adult "coach", who was on friendly terms with the children, asked them to "trick" someone else in a subsequent interview and to pretend that they had played with a certain toy. The children varied in age from 2.5 to eight years. They found that 60% of the children participated in the lie (more older than younger children) and that 35% maintained the ruse throughout. In another coaching study, Bussey (1992) found, as did Tate *et al.* (1992), that older children (five year olds) were more likely to lie than younger children (three year olds). However, 24% of the three year olds were willing to lie when instructed to do so, a percentage that rose to 50% when they were instructed in a very stern manner.

Lewis (1993) describes how children learn to tell lies. Children are encouraged by their parents or care-givers to respond with enthusiasm when receiving a disappointing gift from a relative. Children also notice that others, for example their parents, tell these sort of lies and they then start imitating this behaviour. Lewis also describes the process of how children learn to lie in order to avoid punishment. For example, a two-year-old girl is told by her mother not to eat a biscuit. Later, when the mother asks the girl whether she ate a biscuit, the child admits that she has done so. Her mother then becomes angry and punishes the

[6]The discussion in the literature focuses on the question as to whether these examples of early deception are forms of genuine deception. Are these deliberate attempts by children to get someone to believe something that the children know or believe to be false? Or are these forms of "practical deception" whereby the "liars" want to achieve a goal by saying or doing something that they know or believe to be false without actually trying to affect the belief of the person who is duped? (Sinclair, 1996, see also Lee & Cameron, 2000)? This discussion goes beyond the scope of this book.

child. After only a few interactions like this, the girl learns that if she admits to any wrongdoing, she will be punished. Therefore, she starts to lie in order to avoid punishment. The child sometimes gets away with her lie but at other times her parents detect her lies. In the latter case, the parents will tell the child that it is bad to lie and that she will be punished if she tells a lie. Now the child faces a difficult problem. If she tells the truth about her wrongdoing, she will be punished, and, if she is caught lying about her wrongdoing, she will be punished as well. She, however, already realises that her parents do not detect all her lies. It is therefore better for her to conceal any wrongdoing, and to admit such behaviour only when it is detected.

Personality

Attachment styles

Attachment theory, proposed by Bowlby (1969, 1973, 1980), and expanded on by other researchers, is based on the premise that during childhood individuals develop "internal working models" of self and others based on their experiences of the availability and responsiveness of their parents or care-givers. Children then use these internal working models of relationships to shape their expectancies of other individuals, which in turn influence their interpersonal behaviour, interaction, and typical style of relationships. These internal working models (attachment styles) have been found to be stable over time, and continue to play an active role throughout adulthood.

Attachment styles can be defined according to two dimensions: avoidance and anxiety (Brennan, Clark, & Shaver, 1998; Cole, 2001; Hazan & Shaver, 1987). *Avoidant* attachment is characterised by a negative model of other people (Bartholomew & Horowitz, 1991), lack of trust, fear of intimacy, and avoidance of closeness due to expectations that others will not be available and supportive (Brennan *et al.*, 1998; Cole, 2001; Mikulincer & Shaver, 2005). *Anxious* attachment represents a negative model of self (Bartholomew & Horowitz, 1991), low levels of self-esteem, preoccupation with intimacy, jealousy, longing for closeness, fear of abandonment, and high need for and dependency on other's approval in close relationships rather than an internal sense of self-worth (Brennan *et al.*, 1998; Cole, 2001).

Attachment styles affect how often people lie in romantic relationships (Cole, 2001; Ennis, Vrij, & Chance, in press). *Avoidant* individuals lie more often to their romantic partners than non-avoidant individuals. Avoidant people, who are uncomfortable with intimacy, use deception to keep others at a safe distance (Cole, 2001). Through deception, they protect their privacy, establish the boundary between themselves and their partners, and develop a sense of relational independence. *Anxious* individuals lie more often to their partners than non-anxious individuals. Anxious people, who fear their partners' commitment and availability, use deception to create a false positive image of themselves, an image they believe looks desirable to their partners in order to regulate their partners' interest and devotion to them (Cole, 2001).

Box 2.3 Targets' reactions as a result of attachment style

Jang, Smith, and Levine (2002) examined what happened in romantic relationships when the targets discovered that they were lied to by their partner about a matter of some consequence. The reactions were dependent on the attachment style needs of the target. Those who scored low on avoidance and low on anxiety (these are the people who have the least intent to lie in a relationship) were likely to talk about the issue when they discovered that they have been lied to. They also opted to continue the relationship. Those who scored high on anxiety and low on avoidance talked around and avoided the issue, and also preferred to stay in the relationship. Those who scored high on avoidance and high on anxiety (these are the people most likely to lie in a relationship) were more likely to avoid their partner and not to talk to them at all. They were also more likely to terminate the relationship after deception had been discovered.

Psychopathy

When considering frequent liars, I suspect that many people will think immediately of psychopaths. Psychopathy is a socially relevant personality disorder characterised by callous, remorseless disregard for others (MacNeil & Holden, 2006). Psychopaths enjoy success at the expense of others and are "social predators who charm, manipulate, and ruthlessly plough their way through life, leaving a broad trail of broken hearts, shattered expectations, and empty wallets in their wake" (Hare, 2003, p. xi). Psychopaths use deception to exploit others (Book, Holden, Starzyk, Wasylkiw, & Edwards, 2006; Seto, Khattar, Lalumiere, & Quinsey, 1987), and lie more persistently and blatantly, with considerably more panache than do most other people (Hare, Forth, & Hart, 1989; Porter & Woodworth, 2007).

Machiavellianism and social adroitness

A stereotypical view of liars is that they are selfish, crafty, and manipulative. This is partly true (Hunter, Gerbing, & Boster, 1982; Kashy & DePaulo, 1996; Wilson, Near, & Miller, 1998; see also the information above about psychopaths). People high in *Machiavellianism* view others cynically, show little concern for conventional morality, and openly admit that they will lie, cheat, and manipulate others to get what they want. They also frequently tell self-oriented lies (Kashy & DePaulo, 1996). The term Machiavellianism is derived from the Italian statesman and writer Machiavelli (1469–1527). In 1513 he published his book *Il Principe* (The Prince) in which he made a plea for a united Italy with a powerful ruler to protect the national interest. This can be done by all necessary means, including those that are not morally justifiable. In his book *L'Arte Della Guerra* (The Art of War), published in 1520, he outlined several possibly effective means.

Interpersonally, manipulators are scheming but not stupid. They do not exploit others when their victims might retaliate, and do not cheat when they are likely to get caught. In conversations, they tend to dominate, but they also seem relaxed, talented, and confident. They are usually popular: they are liked more than people who are low in manipulative skills and are preferred as partners (Kashy & DePaulo, 1996).

Social adroitness is also related to manipulativeness but, compared to Machiavellianism, more subtly and with less negative connotations (Kashy & DePaulo, 1996). People high in social adroitness frequently tell self-oriented lies (Kashy & DePaulo, 1996) and tend to persist in lying when challenged (Vrij & Holland, 1998).

Extraverts and introverts

That frequent liars could be popular (e.g., people high in Machiavellianism) becomes clear when taking into account how often sociable people lie. Sociability refers to the tendency to affiliate with others and to prefer being with others rather than being alone. *Extraverts* are sociable people. They like being with other people and are attracted to social life. In contrast, *introverts* are typically more reserved in social contexts. This could be because they prefer their own company and prefer to focus themselves on thoughts and reflections that deal solely with the self. Extraverts lie more often than introverts, even when controlling for the fact that extraverts have more social interactions than introverts (Kashy & DePaulo, 1996; Weiss & Feldman, 2006). Due to their candour, we should label introverts as *honest*, however, instead we often label them as *socially awkward*, perhaps because they tend not to lie in situations where others feel it appropriate to lie.

Public self-consciousness

People differ in how much they care about the impression they make on others. Those who are particularly concerned score high in *public self-consciousness* and *other-directedness*. Such individuals tend to mould themselves to make a good impression on others. One way of doing this is by pretending to be the kind of person they think others may like. Indeed, Kashy and DePaulo (1996) found that such people tell many lies, particularly self-oriented lies. The frequent occurrence of self-oriented lies indicates that the basis of their orientation towards others is a self-centred concern rather than an altruistic concern.

Social anxiety

Expressing true feelings and thoughts sometimes requires self-confidence. Perhaps people who are *socially anxious* – who feel uncomfortable in the presence of others – lack such self-confidence, and therefore lie more often. DePaulo's diary study found no evidence that socially anxious people lie more frequently than socially non-anxious people. However, one could argue that it probably depends on the situation. Socially anxious people may well feel comfortable in some situations, for example, when they are with family or friends. It could thus be that a tendency to lie only arises in situations where they feel uncomfortable, for example when they are with people they do not know well or with authority figures.

In an experiment, we took our participants into an uncomfortable situation (Vrij & Winkel, 1992b). A police officer interviewed the participants about the possession of a set of headphones. All participants were given a small set of headphones prior to the interview (a fact which was unknown to the police officer), and they were all requested to deny this possession. The police officer thoroughly interviewed each participant about the possession of the headphones. In such an uncomfortable situation, social anxiety was related to lying. Fewer socially anxious people (54%) than socially non-anxious people (81%) persisted in their denying throughout the whole interview that they had the set of headphones in their pockets.

Actors

Some people are better *actors* than others and are more skilled in regulating their verbal and nonverbal behaviour than others. The four constructs "emotional control", "social control", "acting", and "social expressivity" are related to acting (Riggio, 1986). Emotional control refers to the ability to regulate emotional communications and nonverbal displays (i.e., the ability to conceal true feelings – for example, being good at maintaining a calm exterior, even when upset). Social control includes role-playing ability, regulation of verbal behaviour, and self-presentational skills. Acting also refers to someone's role-playing ability, while social expressivity includes skill in verbal expression and verbal fluency. These skills benefit liars. In our experiment (Vrij & Holland, 1998), undergraduate students were thoroughly interviewed by a police detective about the course they were studying. Although none of the participants were, they were all asked to pretend that they were studying psychology. The police detective asked questions such as: "Can you tell me what topics you have recently studied?", "Can you tell me something about these topics in more detail?", "Tell me about your coursework titles?", "What do your exams cover?", and "Name three famous psychologists?" Compared to non-actors, actors persisted more in lying when challenged.

LYING AS A SOCIAL LUBRICANT AND SELFISH ACT

This chapter challenged the conventional view that lying is necessarily bad. Conversations could become awkward and unnecessarily rude and social interactions could easily become disrupted when people tell each other the truth all the time. Introverts are relatively honest, but they are considered by others to be somewhat socially awkward, possibly because they don't tend to lie at times that others feel it is appropriate to lie. In contrast, extraverts lie frequently, and are considered to be socially skilled. Machiavellians are also frequent liars and are often popular. In fact, if asked to make a likeable impression, people often choose to lie. In their experiment, Feldman *et al.* (2002) asked some participants to present themselves so that their conversation partners would think that they are likeable; other participants were not given a specific goal. Those who were asked to appear likeable lied more in the subsequent interaction than the other participants.

We tell lies even to people we feel emotionally close to: we tell many lies at the beginning of a romantic relationship and make many untruthful or exaggerated flattering remarks to people we like. Obviously, when we lie to our romantic partners about serious issues and we are caught out, we may be in trouble. Our partners may even terminate the relationship. However, we saw in this chapter that it is often not the act of lying that make partners decide to end a relationship, but the information that the liar attempted to hide from them.

The nature of lying is two-pronged, and how we feel about deception depends on the reason why the lie is told (Seiter *et al.*, 2002). Most lies are told for psychological reasons and people don't feel bad about telling such lies. We do not relish having to express all our thoughts (e.g., "I find that woman more attractive than my own partner") and would rather lie. Neither do we always wish to show our true self. We prefer to edit the self somewhat so that we present ourselves in a way we would like to be seen by others. Psychological lies are often told to protect ourselves or to avoid tension and conflict in social interactions and to minimise hurt feelings and ill-will (DePaulo, Kashy *et al.*, 1996). They thus sometimes act as a social lubricant and improve social relationships. In this context unsurprisingly, 70% of the participants in DePaulo's diary study informed the researchers that they would tell the lie again. This could also explain why Seiter *et al.* (2002) and Arnett Jensen *et al.* (2004) found that people thought that it is sometimes highly acceptable to lie to others they feel close to, and even more acceptable than to lie to people they do not feel so close to. This is particularly the case when the motives of these lies are to avoid conflict or to benefit others (see also Barnett, Bartel, Burns, Sanborn, Christensen, & White, 2000; Pontari & Schenkler, 2006). I also discussed another reason why people sometimes lie to those who feel emotionally close to: they do so to maintain their privacy.

This positive aspect of deception also explains why we sometimes don't mind being lied to. We would not appreciate it if others are always perfectly honest with us, and if our friends always told us their real thoughts about us and everything we do. What would happen to our self-esteem if we always knew what people really think of us? Given that lies may serve as a social lubricant, being ignorant about the truth often serves us well. It is in not in our interest to be capable of detecting all lies, and we sometimes do not even attempt to detect lies (I labelled this the *ostrich effect* in Chapter 1).

However, sometimes the situation is different, and most people will come across situations where they really would like to know the truth. The viewer wants to know whether the politician is telling the truth when she denies any involvement in a bribery scandal; the general wants to know whether he can trust his opponent before he signs the truce agreement; the teacher wants to know which pupil played a trick upon him; the mother wants to know whether her daughter really has finished her homework; the possible purchaser wants to know whether the second-hand car is really as good as the salesperson suggests; the employer wants to know whether the candidate is indeed as capable as the candidate says; the customs officer wants to know whether the traveller really has nothing to declare; and the police detective wants to know whether the suspect's alibi is reliable. Being able to detect these sorts of lies would benefit individuals or the society as a whole. The remaining part of the book deals with how liars respond and how they possibly could be detected.

CHAPTER 3

Nonverbal Behaviour and Deception

> He who has eyes to see and ears to hear may convince himself that no mortal can keep a secret. If his lips are silent, he chatters with his finger-tips; betrayal oozes out of him at every pore (Freud, 1959, p. 94).

Freud's citation may suggest that diagnostic nonverbal cues to deception exist. From fiction comes the cue of Pinocchio's growing nose. His nose was unaltered each time he spoke the truth, but grew larger each time he lied. In this chapter I summarise three theoretical perspectives discussing why differences in nonverbal behaviour may occur between truth tellers and liars. All three perspectives have two important features in common: they all predict that the mere fact that people lie will not automatically change their behaviour; and none of them rule out that cues that may occur when people lie could also occur when they tell the truth. In other words, no theoretical perspective predicts that diagnostic nonverbal cues to deception, akin to Pinocchio's growing nose, exist.

I then evaluate the published studies on nonverbal cues to deceit in light of these theories. Indeed, these studies reveal that a sign equivalent to Pinocchio's growing nose has not been found. In fact, most of the nonverbal cues described in this chapter do not appear to be related to deception. However, for some cues a weak relationship with deception emerged. I discuss possible reasons why only a few and rather weak relationships between nonverbal behaviour and deception have been found. Could it be that more nonverbal cues to deception exist than we are aware of, but that researchers so far have overlooked some diagnostic nonverbal cues to deception? And could it be that a combination of nonverbal cues, rather than individual cues, reveal diagnostic indicators of deceit? I further address the possibility that each individual does show clear nonverbal signs of deceit, but that different individuals show different signs. This would result in no signs of deceit emerging when these individuals are analysed at a group level (what typically happens in deception studies), but a different picture would emerge if each individual is examined separately.

I also discuss situational influences on nonverbal cues to deception. For example, in most deception studies the liars were college students who lied for the sake of the experiment. Maybe liars don't show nonverbal signs of deception in those situations, but perhaps they show clear cues under more challenging circumstances. I discuss the behaviours shown by two politicians (Bill Clinton and Saddam Hussein) in televised interviews, and the behaviours displayed by suspects during their police interviews. I finally elaborate on whether the interview styles used by the interviewer can affect the nonverbal cues of deceit shown by the interviewee.

THREE THEORETICAL PERSPECTIVES ON NONVERBAL CUES TO DECEPTION

Zuckerman, DePaulo, and Rosenthal's Multi-Factor Model

According to Zuckerman, DePaulo, and Rosenthal (1981) the presence of three factors could influence cues to deception: (1) emotional reactions; (2) cognitive effort; and (3) attempted behavioural control.[1] Each of these factors may influence a liar's nonverbal behaviour and emphasises a different aspect of deception, and lies may well feature all three factors.

The emotional approach

Telling a lie is most commonly associated with three different emotions: guilt, fear, and delight (Ekman, 1985/2001, 1989). Suppose that a politician has secretly accepted a large sum of money from a company

[1] Zuckerman *et al.* (1981a) mentioned a fourth factor "arousal". However, they acknowledge that arousal overlaps with the emotion factor (Zuckerman *et al.*, 1981a, p. 9).

in exchange for lobbying, and that a journalist becomes suspicious and asks the politician about his links with the company. While denying any wrongdoing, the politician may feel *guilty* because he realises that it is wrong to deceive the journalist.[2] He also may be *afraid*, because he could be worried that the journalist will find out that he is lying, which would probably result in the end of his political career. Or he may become very *excited* because he enjoys the opportunity of fooling the journalist, or enjoys the moment when he realises that his lie was successful.

Guilt, fear, and excitement can influence a liar's behaviour in several ways. For example, guilt may result in gaze aversion because the liar does not dare to look the target straight in the eye while telling a barefaced lie. Fear may result in increased physiological arousal, and this could lead to an increase in cues such as eye blinks, self-adaptors (touching own clothes, hair, face, etc.), speech hesitations (mm's and er's), speech errors (stutters, repetition of words, omission of words), and fundamental frequency of pitch (i.e., higher pitched voice).

Experiencing negative emotions (e.g., guilt and fear) may also result in signs of withdrawal such as less eye contact, less body orientation (degree to which body and head are oriented to the other person), and a decrease in illustrators (gestures people make that accompany their speech). Excitement may result in behavioural signs of joy, such as an increase in movements and in smiling.

The cognitive effort approach

Lying sometimes requires extra mental effort, and several aspects of lying contribute to this increased mental load.[3] First, formulating the lie itself may be cognitively demanding. Liars may need to make up their stories and must monitor their fabrications so that they are plausible and adhere to everything the observer knows or might find out. In addition, liars must remember their earlier statements, so that they appear consistent when retelling their story, and know what they told to whom. Liars should also avoid making slips of the tongue, and should refrain from providing new leads.

[2] Obviously, he may also feel guilty because he acknowledges that accepting a bribe is morally wrong.

[3] Lying is not always difficult, and sometimes it is even easier to lie than to tell the truth (McCornack, 1997). Suppose a friend gives you a present for your birthday that you don't like. In this case it is probably easier to pretend that you like the present than to say that you don't like it.

A second aspect of lying that adds to mental load is that liars are typically less likely than truth tellers to take their credibility for granted (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003; Kassin, 2005; Kassin & Gudjonsson, 2004; Kassin & Norwick, 2004). There are at least two reasons for this. The stakes (i.e., negative consequences of getting caught and positive consequences of getting away with the lie) are sometimes higher for liars than for truth tellers. Smugglers are probably keener to make an honest impression on customs officers than non-smugglers, because the negative consequences for having to open their suitcases are much higher for smugglers than for non-smugglers. In addition, truth tellers typically assume that their innocence shines through (Granhag, Strömwall, & Hartwig, 2007; Kassin, 2005; Kassin & Gudjonsson, 2004; Kassin & Norwick, 2004; Vrij, Mann, & Fisher, 2006b), which could be explained with the *illusion of transparency* (Gilovich, Savitsky, & Medvec, 1998), the belief that "one's inner feelings will manifest themselves on the outside", and the *belief in a just world* (Lerner, 1980), the belief that "people will get what they deserve, and deserve what they get". As such, liars will be more inclined than truth tellers to monitor and control their demeanour so that they will appear honest to the lie detector (DePaulo & Kirkendol, 1989; this is emphasised in the attempted control approach below). Paying special attention to their own demeanour could be cognitively demanding for liars.

Third, because liars do not take their credibility for granted, they may be inclined to monitor the interviewer's reactions more carefully in order to assess whether they are getting away with their lie (Buller & Burgoon, 1996; Schweitzer, Brodt, & Croson, 2002). Carefully monitoring the lie detector also requires cognitive resources. Fourth, liars may be pre-occupied by the task of reminding themselves to act and role-play (DePaulo, Lindsay *et al.*, 2003), which requires extra cognitive effort. Fifth, liars have to suppress the truth while they are lying and this is also cognitively demanding (Spence *et al.*, 2001). Finally, while the truth often comes to mind automatically, activation of the lie is more intentional and deliberate, and thus requires mental effort (Gilbert, 1991; Walczyk, Roper, Seemann, & Humphrey, 2003; Walczyk, Schwartz, Clifton, Adams, Wei, & Zha, 2005).

Evidence has suggested that people engaged in cognitively complex tasks blink less (Bagley & Manelis, 1979), make more speech hesitations and speech errors, speak slower, pause more, and wait longer before giving an answer (Goldman-Eisler, 1968). Cognitive complexity also leads to fewer hand and arm movements and to more gaze aversion (Ekman, 1997; Ekman & Friesen, 1972). This decrease in hand and arm movements occurs because cognitive demand results in a neglect of body language, reducing overall animation. For example, research examining truth tellers' and liars' brain activity showed that deception was related to increased activity in "higher" areas of the brain (Spence *et al.*, 2004), which, in turn, inhibits fidgety movements (Shallice & Burgess, 1994). Gaze aversion (usually to a motionless point) occurs because looking the conversation partner in the eye can be distracting (Doherty-Sneddon, Bruce, Bonner, Longbotham, & Doyle, 2002; Doherty-Sneddon & Phelps, 2005).

The attempted behavioural control approach

Liars may realise that observers pay attention to their behavioural reactions to judge whether they are lying. Liars therefore may attempt to control their behaviour (Buller & Burgoon, 1996; Burgoon & Buller, 1994; Burgoon, Buller, Floyd, & Grandpre, 1996; Burgoon, Buller, White, Afifi, & Buslig, 1999; Krauss, 1981). In particular, they may wish to avoid exhibiting behaviours that they believe make a dishonest impression on others and may try instead to show behaviours that they believe appear credible (Hocking & Leathers, 1980; Leary & Kowalski, 1990). However, this is not easy. Some behaviours are almost beyond control of the liar, because they are linked to strongly felt emotions or high stress (Ekman, 1985/2001). Anger, for example, results in several cues, including narrowing of the lips. Most people, however, cannot narrow their lips voluntarily. Thus, if an angry person denies being angry, an automatic response such as the narrowing of his lips may give the lie away. Similarly, tone of voice is difficult for senders to control (Ekman, 1981), because the automatic nervous system controls the relevant features of the voice in moments of high stress (Hocking & Leathers, 1980).

A further difficulty in controlling behaviour is that people cannot be silent nonverbally. Suppose a guilty suspect realises during his police interview that the police know more than he thought they did about his involvement in the crime. This may throw the suspect off guard and he may need time to recompose himself. Police officers can observe and interpret this behaviour.

Another difficulty in controlling behaviour is that people are often not aware of their behaviour in ordinary day-to-day situations. There are two reasons for this lack of awareness. First, unless we look in the mirror or watch a videotape of ourselves, we don't actually see ourselves. Second, we exchange information predominantly via words. When we are asked to describe our activities of that day, we mostly choose to use words to describe what we have done.[4] We therefore concentrate more on our speech and neglect our nonverbal behaviour somewhat. Lack of insight into our own behaviour may sometimes make us unaware of subtle changes that take place in our behaviour. For example, a suspect in a police interview will be aware of what he says when providing his alibi, but he may not realise that he speaks somewhat slower when providing this alibi; and an applicant in a selection interview will be aware what she says about her relationship with her current employer, but she may not notice that she looks grumpy when she mentions his name. As long as people are unaware of changes in their behaviour, they are unlikely to address these changes, and, subsequently, the changes may leak valuable information to observers.

Even if people actively address their behaviour and try to appear credible, it is not guaranteed that they will manage to do so. The attempted behavioural control approach predicts that those behaviours that are the most difficult to control will give the lie away (Ekman & Friesen, 1974). According to this reasoning, gaze is unlikely to be an indicator of deceit. People can make clear via their gaze whether they are interested in someone's conversation and whether they want to say something (Ekman, 1985/2001; Kleinke, 1986). Also, people attempt to persuade others by looking them in the eyes (Kleinke, 1986). The great communicative potential of gaze means that people are practised at using and therefore controlling their gaze. Since the common view is that liars look away,[5] liars will therefore probably attempt to avoid gaze aversion, and they are likely to be reasonably successful in doing this.

In contrast, all sort of movements people make are not salient in communication and are less often attended to and reacted to by others. We are therefore less practised in controlling or deliberately making movements, tending to display them unintentionally. Since the common view is that liars increase their movements, liars will probably move very deliberately and tend to avoid those movements that are not strictly essential. This will result in an unusual degree of rigidity and inhibition, and in behaviour that will appear planned, rehearsed and lacking in spontaneity, because people normally make movements with their body, hands, fingers, feet, legs, and so on, that are not essential (Burgoon & Buller, 1994; DePaulo & Kirkendol, 1989).

[4] There are exceptions. For example, it is sometimes easier to describe the shape and size of objects (the shape of a lamp, the size of a CD) and the location of an object (the ball is there) with hand and arm movements rather than with words.

[5] I will discuss in Chapter 5 which behaviours people typically associate with deception.

Like most movements, vocal characteristics such as speech hesitations, speech errors, and pauses in speech are usually made unintentionally and are not usually important in the exchange of information. We therefore may assume that people do not often practise controlling these behaviours, and so are not very good at controlling them. Since the common view is that liars tend more often to make speech hesitations, errors and pauses, liars will probably try to avoid making such nonfluencies. This, however, may result in a speech pattern that sounds unusually smooth, as it is normal for most people to make some hesitations, errors, and pauses in speech.

Finally, as I already mentioned above, some nonverbal cues such as facial expressions of emotions or pitch of voice are difficult to control when emotions are strongly felt or the liar experiences high stress. Such nonverbal cues may betray a liar.

Another possible cue that may result from inadequate control of behaviour is that performances may look flat, because the liar lacks involvement in his claims (Burgoon & Buller, 1994; Burgoon, Buller, Dillman, & Walther, 1995; DePaulo, Lindsay *et al.*, 2003). An artist who applies for a job as salesperson because he needs the money may not look convincingly enthusiastic during the selection interview; and a mother who punishes her child for wrongdoing may not look sincere if she, in fact, was amused by the child's behaviour. Charles Ingram, who was found guilty in the United Kingdom of cheating his way to the top prize in the popular TV quiz *Who wants to be a Millionaire* may also have experienced this apparent indifference. Staff working for the TV programme became suspicious when Ingram and his wife, after winning the top prize of £1,000,000, "had not appeared as jubilant as the newly rich might" (*The Independent*, 8 April 2003, p. 9).

Factors influencing the three processes

Whether, and to what extent, liars actually experience emotions or cognitive load, or try to control their behaviour depends on the personality of the liar and the circumstances under which the lie takes place.

Guilt Regarding personality, some people will experience less guilt than others while lying. For people high in Machiavellianism or social adroitness, lying is a normal and acceptable way of achieving their goals (Chapter 2), and, consequently, they typically don't feel uncomfortable or guilty when they fool others (Kashy & DePaulo, 1996; Gozna, Vrij, & Bull, 2001). There is also evidence that women and men cope differently: women become somewhat more uncomfortable while telling lies than do men (DePaulo, Epstein, & Wyer, 1993; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996). DePaulo and Kirkendol (described in DePaulo, Epstein, and Wyer, 1993) asked men and women about the most serious lie they ever told to anyone, and the most serious lie that anyone ever told them. These serious lies had a bigger impact on women than on men. When discussing the situations in which they were the ones who told a serious lie, women described themselves as more guilty and tearful than men.

Experiencing guilt is also related to the circumstances under which the lie takes place. Telling high-stakes lies (lies that are associated with large positive consequences of getting away with the lie or large negative consequences of getting caught) probably result in experiencing more guilt than telling low-stakes lies. For example, the restaurant patrons who deliberately do not mention to the waiter that they have not been charged for their drinks would probably feel more guilty had they been drinking champagne than had they been drinking water.

A liar will not feel guilty when the lie can be morally justified. A spy, for instance, tries to protect the national interests of his or her country and therefore finds it entirely acceptable to lie in order to do so, and few people experienced moral objections when lying to German soldiers when their countries were occupied in World War II (Ekman, 1985/2001). Neither will a liar feel guilty when he or she believes that it is legitimate to lie. A salesman considers it to be part of his job to exaggerate the favourable aspects of the product he is trying to sell and will therefore not experience guilt when doing so. A woman who is having an adulterous affair may not feel guilty about it if she believes that her husband is also cheating on her. Intriguingly, the latter belief (that the other is also dishonest) is a common thought amongst liars (Sagarin, Rhoads, & Cialdini, 1998).

Fear The amount of fear experienced during deception depends on how good liars think they are at lying. Some people are very good liars and they know it: by experience they have learned that it is easy for them to fool others and that they hardly ever get caught. This will increase their self -confidence and will decrease their detection anxiety while lying.

Since the outcomes of high-stakes lies really matter to liars, they will probably experience more fear of getting caught when telling high-stakes lies compared to when telling low-stakes lies. Also, the negative consequences of getting caught are typically higher when people lie out of self-interest than when they lie to help others, and this may result in experiencing more fear when telling self-oriented lies compared to when telling other-oriented lies (Anderson, Ansfield, & DePaulo, 1999). The amount of detection anxiety liars experience further depends on the person to whom the lie is told. When a liar thinks that the target is very good at detecting lies, the liar will experience more fear than when he or she thinks it is easy to dupe that person.

Excitement The excitement that accompanies lying, so-called duping delight (Ekman, 1985/2001), will increase when the person to whom the lie is told is known to be difficult to deceive. The excitement will also increase when there are onlookers when the duping takes place. The girl who tries to fool her teacher will be more excited when other pupils are in the classroom than when she is alone with the teacher.

Cognitive effort The extent to which lying is cognitively demanding depends on people's personality. People high in Machiavellianism don't find lying too cognitively complicated (Kashy & DePaulo, 1996; Gozna *et al*., 2001). Moreover, verbally skilled people find it easier to lie (Kashy & DePaulo, 1996; Vrij, Akehurst, Soukara, & Bull, 2002, 2004a), perhaps because they find fabricating stories relatively undemanding. People who find themselves good at acting also find lying easier (Gozna *et al*., 2001; Vrij, Edward, & Bull, 2001c), perhaps because they find controlling their behaviour relatively easy. The amount of cognitive load experienced when lying further depends on a person's intelligence. Less intelligent people find it harder to lie (Ekman & Frank, 1993; Vrij & Mann, 2001a), perhaps because they have more difficulty in keeping their stories straight, particularly when challenged. It may also be that the multiple tasking of telling a plausible story, monitoring their behaviour, and observing the lie detector may be more mentally difficult for them than for more intelligent people.

Since the outcomes of high-stakes lies really matter to liars, they probably try particularly hard in such situations not to get caught. This increased effort will be cognitively demanding and therefore liars probably experience more cognitive load when telling high-stakes lies compared to when they tell low-stakes lies. The higher negative consequences that are typically associated with self-oriented lies may also result in experiencing more cognitive demand when telling self-oriented lies than when telling other-oriented lies. Telling an outright lie may be more cognitively demanding than concealing information, because when liars conceal information they don't have to invent a story or remember what they have said. Moreover, telling an elaborate lie is more cognitively demanding than providing short "yes" and "no" answers, because there is more information to fabricate and to remember when telling elaborate lies (Vrij, Mann, & Fisher, 2006b). Lying is probably also more demanding when the lie is not well prepared or rehearsed in advance.

Attempted behavioural control Some people are more motivated than others to make a good impression. For example, people differ in the extent to which they think others scrutinise them. People who score high in Public Self-Consciousness particularly see themselves as the focus of other people's attention (Fenigstein, Scheier, & Buss, 1975). Due to this heightened awareness of being the focus of attention, people high in Public Self-Consciousness are particularly motivated to make a good impression on others and are particularly likely to attempt to control their behaviour to make an honest impression (Vrij, Edward, & Bull, 2001c).

Also, the more people think that the impression they make is important in accomplishing their goals, the more motivated they are to engage in impression management (Leary & Kowalski, 1990). Therefore, people high in Machiavellianism are particularly likely to engage in strategic self-presentation to influence others, because they see the impression they create as relevant to accomplishing their goals (Leary & Kowalski, 1990). Finally, it sounds reasonable that liars will be more concerned with the impression they make in some situations than in others. For example, they will try harder to control themselves when getting away with the lie really matters to them (e.g., high-stakes lies) compared to when the lie is trivial (e.g., low-stakes lies).

DePaulo's Self-Presentational Perspective

Zuckerman *et al.*'s (1981a) perspective predicts that the more liars experience one or more of the three factors (emotion, cognitive effort, attempted behavioural control), the more likely it is that cues to deception will occur. A key element in Bella DePaulo's self-presentational perspective (DePaulo 1992; DePaulo, Lindsay *et al.*, 2003) is that truth tellers and liars have much in common. She points out that, for example, truth tellers may also experience emotions and cognitive load, or may try to control themselves, and that this may also influence truth tellers' behaviour. Thus, guilty (deceptive) suspects may be afraid of not being believed during their police interviews, but so will innocent (truth telling) suspects, because they too could face negative consequences if they are not be believed (Ofshe & Leo, 1997). Because of that fear, they may show the same nervous reactions as guilty suspects who are afraid of being caught (Bond & Fahey, 1987).

According to DePaulo, Lindsay *et al.* (2003), truth tellers and liars will only succeed in their social interaction goals if they appear sincere. The difference between truth telling and lying is that the liar's claim to honesty is illegitimate, and this lack of legitimacy has two implications. First, deceptive self-presentations may be less embraced than truthful self-presentations. Liars may endorse their statements less because they have moral scruples, lack emotional investment in their false claims, or lack the knowledge and experience to back up their deceptive statements convincingly. Second, liars typically experience a greater sense of awareness and deliberateness in their performances than truth tellers, because they may take their credibility less for granted than truth tellers. As already explained, they do this because the stakes are sometimes higher for liars than for truth tellers but also because truth tellers believe that their innocence will shine through. Although truth tellers are also keen to be seen as truthful, they typically do not think that this will require any special effort or attention. Taking credibility less for granted implies that liars are more likely than truth tellers to think that it is important to make a convincing impression on others, and, subsequently, liars are more concerned about the impression they make on others than truth tellers.

Box 3.1 Suspicious but innocent

Ironically, when truth tellers are not concerned about the impression they make on others, they could end up showing behaviour that appears suspicious. The following real-life example illustrates this (Vrij & Mann, 2001b). A man gave a press conference that was broadcast on British television in which he asked the public for information about the whereabouts of his missing daughter. In his broadcast the man said: "We want you home. You may just have it in your head (long pause) to find it difficult to come home for whatever reason. Don't find it difficult to come home, please come home". Meanwhile the man showed remarkable behaviour. He had "shifty" eyes and appeared to be smiling. During the pause in his statement he touched his mouth with his hand. When I show this fragment during workshops, most observers find his behaviour suspicious. Indeed, as we will see in Chapter 5, this is the type of behaviour many people do find suspicious, and police manuals often mention that liars show this type of behaviour. However, the man was not a suspect in this case. In fact, I believe that it is unlikely that the man would have exhibited this behaviour if he had been lying, because if he was then he would probably have realised that his behaviour evokes suspicion and so he would have refrained from displaying it. His innocence, however, made him less aware of the impression he made on others, resulting in the behaviour he showed.

Buller and Burgoon's Interpersonal Deception Theory (IDT)

A third perspective on deception, Buller and Burgoon's (1996) Interpersonal Deception Theory (IDT) postulates that during face-to-face encounters, liars must accomplish numerous communication tasks simultaneously. They must produce a credible verbal message and display credible nonverbal behaviour simultaneously. They must also attend to their conversation partner to find out whether they are still believed, while managing their emotions, keep dialogue running smoothly, responding appropriately to what the conversation partner says, and be discreet about any intentions to deceive (Buller & Burgoon, 1996). IDT embraces Zuckerman *et al.*'s (1981a) factors (emotion, cognitive effort, and attempted behaviour control) as underlying reasons for cues to deceit (Burgoon *et al.*, 1999). In addition, it emphasises that, when deception occurs in interactive contexts, it is not a unidirectional activity. Rather, both liar and target mutually influence each other (Burgoon, Buller, Floyd *et al.*, 1996).

According to IDT, targets' behaviour may influence senders' behavioural displays both directly, via synchrony, and indirectly, because it may trigger behavioural adjustments (Burgoon *et al.*, 1999). Regarding the direct effects, when people communicate with each other, matching and synchrony may take place (Burgoon, Buller, Dillman, & Walther, 1995; Burgoon, Buller, Ebesu, White, & Rockwell, 1996; Burgoon *et al.*, 1999; Chartrand & Bargh, 1999). People may mirror each other's posture, or they may converge in how quickly and how loudly they speak. They may also reciprocate each other's gazing, nodding, accents, movements, smiling behaviour, and facial expressions (Akehurst & Vrij, 1999; Baumeister, Hutton, & Tice, 1989; Cappella & Schreiber, 2006; DePaulo & Friedman, 1998; Dimberg, Thunberg, & Grunedal, 2002; Tickle-Degnen, 2006). This *chameleon effect* (Chartrand & Bargh, 1999) emerges even when strangers interact with each other, and it happens typically within a few minutes (Chartrand & Bargh, 1999).

The indirect effects are related to feedback from the target. Sometimes liars' statements are received with scepticism. For example, the journalist does not believe the police chief when she says that she is not aware of any unlawful practices in her force, and the judge distrusts the hooligan when he claims that he only used his knife to threaten the fellow supporter and not to stab him. When liars are exposed to negative feedback from the target, expressed through either verbal comments or through nonverbal behaviour, liars may realise that their performance is lacking credulity. Consequently, liars may respond by making behavioural adjustments to diminish suspicions and to show "honest behaviour", such as looking the target straight into the eyes, avoiding fidgeting and so on (Buller, Comstock, Aune, & Strzyzewski, 1989; Buller, Strzyzewski, & Comstock, 1991; Stiff & Miller, 1986).

Summary

The three theoretical perspectives discussed here make clear that the relationship between lying and nonverbal behaviour is complex. Zuckerman *et al.*'s (1981a) assumptions that liars may show signs of emotions and cognitive load seem straightforward. However, those assumptions may lead to opposite behaviours. For example, arousal typically leads to an increase in eye blinks whereas cognitive load typically leads to a decrease in eye blinks; and the emotional approach predicts an increase in certain movements (signs of nervous behaviour), whereas the cognitive load approach predicts a decrease in movements during deception as a result of neglecting the use of body language. DePaulo, Lindsay *et al.*'s (2003) self-presentation perspective stresses that experiencing emotions or cognitive effort is not the exclusive domain of liars. Truth tellers may also experience these and, as a result, may also display nonverbal cues associated with emotion or cognitive load.

Zuckerman *et al.*'s (1981a) attempted behavioural control prediction is not straightforward because the behaviours shown by deceptive senders, as a result of this deliberate control, will depend on not only their perceptions of what constitutes a credible nonverbal display but also their acting skills in performing this display. However, some general predictions can be made on the basis of the attempted behavioural control approach that contradict what can be expected on the basis of the emotional and cognitive effort approaches. The emotional and cognitive effort approaches predict an increase in speech hesitations and speech errors due to nervousness and cognitive load respectively, whereas the attempted control approach predicts that liars will attempt to suppress most of those speech disturbances, and will therefore sound atypically smoothly. The emotional and cognitive effort approaches predict an increase in gaze aversion as a result of nervousness and cognitive load, whereas the attempted control approach predicts that liars will be as good as truth tellers at controlling their gaze behaviour.

Finally, Buller and Burgoon's (1996) IDT's interactive approach implies that deceptive behaviour may be influenced directly by the behaviour of the target (a result of the chameleon effect) or indirectly influenced by the suspicions raised by the target. The complex relationship between nonverbal communication and deception makes it unlikely that clear, diagnostic, nonverbal cues to deception exist.

Deception research, summarised in the next sections, has supported this view.

METHODS OF INVESTIGATING NONVERBAL BEHAVIOUR WHEN LYING

In research where nonverbal cues to deception have been examined, trained raters watch video footage (or sometimes listen to audiotapes) of truth tellers and liars. They analyse with particular coding systems the frequency of occurrence or duration of certain nonverbal behaviours displayed by the truth tellers and liars, and compare the truthful and deceptive responses. In real-life studies, typically called "field studies", video footage of real-life settings is analysed, such as police–suspect interviews (Mann, Vrij, & Bull, 2002). In laboratory studies video footage is analysed of participants who were instructed by researchers to tell the truth or to lie for the sake of the experiment. Field studies probably have greater appeal, due to their realistic nature. However, there are three problems associated with conducting field studies: *obtaining video footage*; establishing the *ground truth*; and selecting *comparable truths*. Regarding obtaining video footage, whilst it may be interesting to examine how a woman responds when she lies to her husband about her affair, or how a teenager responds when he lies to his mother about having taken cash from her purse, it is highly unlikely that such interactions are videotaped. People typically don't have cameras installed in their homes, so many lies told in daily life can therefore not be analysed.

In cases where video footage is available, only those responses can be analysed which are known to be true or false. To establish this so-called ground truth satisfactorily, independent case facts, such as medical evidence, material evidence, DNA evidence, or reliable eyewitnesses are needed. Unfortunately, independent case facts are often not available. Take for example Bill Clinton when he testified before the Grand Jury about his alleged affair with Monica Lewinsky. Was he lying during his testimonial? And if so, exactly when during that interview? We simply don't know, because reliable independent case facts are not available. In other words, in many real-life settings the ground truth cannot be satisfactorily established.

Finally, only truthful and deceptive responses should be analysed which are truly comparable in all aspects other than the veracity status. Only then can differences in behaviour between the truthful and deceptive fragments be attributed to the fact that the person is lying. Selecting comparable truths and lies is difficult, because people behave differently in different situations (Dente, Barath, Ng, Vrij, Mann, &

Bull, 2006; DePaulo & Friedman, 1998; Vrij, 2004b, 2006a). I label those differences *intrapersonal differences*. First of all, the setting is important: (i) people react differently in formal settings, such as during a selection interview, than in informal settings, such as at home with the family; (ii) they also react differently when they are accused of wrongdoing than when they are unchallenged (Vrij, 2006a); and (iii) they show different behaviours when they are interviewed by different people (Vrij & Winkel, 1991). Moreover, behaviour is topic-related. People respond differently when discussing a topic that embarrasses them than when discussing a neutral topic (Kleinke, 1986), and they respond differently when they discuss a topic they care about or is important to them than when they discuss a topic with which they have less personal involvement (Davis & Hadiks, 1995; Matarazzo, Wiens, Jackson, & Manaugh, 1970). Finally, people's behaviour sometimes changes over time in the same interview (Buller & Burgoon, 1996; Burgoon *et al.*, 1999; Stiff, Corman, Krizek, & Snider, 1994; White & Burgoon, 2001), or, if they are interviewed on more than one occasion, changes may occur over repeated interviews (Granhag & Strömwall, 2002). Therefore, when researchers wish to compare a person's deceptive nonverbal response with a truthful nonverbal response from the same person, they need to make sure that the deceptive and truthful responses are taken from the same interview setting; that the person talks about similar topics in the deceptive and truthful parts; and that these parts were discussed within a short period of time from each other.

As a result of the problems with obtaining video footage, field studies about nonverbal cues to deception are sparse. Also, the deception field studies that have been published are often of poor quality, because the researchers were not able to establish the ground truth satisfactorily, failed to select comparable truths, or due to a combination of these two problems.

An alternative to field studies are laboratory studies. In such studies, researchers ask participants, mostly university (college) students,[6] to tell the truth or lie, and measure their nonverbal responses during truth telling and lying. In the studies published to date, participants have told the truth or lied about many different topics. For example, about a film they just had seen, whether they had a certain object in their pocket, or their involvement in money that went missing. In other studies they told the truth or lied about playing cards they were shown, the number of dots that appeared on a screen, their feelings about certain people, or some controversial issues.

[6]See Klaver, Lee, and Hart (2007) and Porter, Doucette, Woodworth, Earle, & MacNeil (in press) for two deception laboratory studies with offenders as participants.

Laboratory studies have some advantages. The participants can be videotaped enabling the researcher to play back their behaviour as many times as desired. Establishing the ground truth is not a problem, as the researchers, who instructed participants to tell the truth or lie, know who is lying. Creating comparable truths is not an issue either, as the situation for truth tellers and liars is identical except from lying. (That is, truth tellers and liars talk about the same topics, are interviewed by the same interviewer, etc.; the only difference is that one tells the truth and the other fabricates.) Differences in responses between truth tellers and liars can therefore be attributed to the fact that one group of participants is lying. Moreover, the controlled nature of laboratory settings enables researchers to examine the influence of all sorts of factors on deceptive behaviour. For example, by asking children and adults to tell the truth or to lie about the same topic, it could be measured whether children show different nonverbal cues to deception than adults (Vrij, Akehurst, *et al.*, 2004a); and by instructing participants of different ethnic origin to tell the truth or to lie about the same topic, it could be examined whether participants with different ethnic background or culture show the same or different nonverbal cues to deception (Vrij & Winkel, 1991). In a similar vein, it allows researchers to investigate whether the interview style used by the interviewer has an effect on the nonverbal cues to deception the interviewees display (Vrij, 2006a); and whether lies that are planned and well rehearsed result in different nonverbal cues than lies than are told spontaneously (Littlepage & Pineault, 1985).

However, laboratory studies do have limitations. In such studies participants do not choose to lie, but are instructed to do so by the experimenter. This means that lying is condoned, and that feelings of guilt are unlikely to occur. Another restriction is that the stakes are never really high (Ekman, 1985/2001; Malone & DePaulo, 2001; Miller & Stiff, 1993). In order to raise the stakes in laboratory experiments, participants have been offered money if they successfully get away with their lies (Vrij, Akehurst *et al.*, 2002; Vrij, Edward, & Bull, 2001c). In other studies, participants are told that they will be observed by a (participant's) peer who will judge their sincerity (DePaulo, Stone, & Lassiter, 1985), or that being a good liar is an important indicator of being successful in a future career (DePaulo, Lanier, & Davis, 1983).[7] Such studies result in rather low-stakes lies. As such they provide useful examples of how people behave when they lie in daily life, because most

[7] The latter is true. For example, good nurses are good liars, because they have the ability to conceal negative emotions when they interact with patients who are terminally ill, or with patients with severe disfigurements (e.g., burns) and so on (Ekman, 1985/2001).

of the lies people tell are low-stakes lies (DePaulo, Kashy *et al.*, 1996, and see below).

However, suspects in police interviews, smugglers at airports, corrupt politicians in conversations with suspicious journalists, and husbands who cheat on their wives, tell high-stakes lies. In an attempt to create examples of such lies, some researchers have raised the stakes further in laboratory studies. For example, participants in Frank and Ekman's (1997) experiment were given the opportunity to "steal" US \$50. If they could convince the interviewer that they had not taken the money, they could keep all of it. If they took the money and the interviewer judged them as lying, they had to give the US \$50 back and also lost their US \$10 per hour participation fee. Moreover, some participants faced an additional punishment if they were found to be lying. They were told that they would have to sit on a cold metal chair inside a cramped darkened room labelled ominously XXX, where they would have to endure anything from 10 to 40 randomly sequenced 110-decibel starting blasts of white noise over the course of one hour.

A study like this raises ethical concerns. Yet, even despite the ethical issue, one might argue that the stakes in such a study do not compete with the stakes in some real-life situations. Providing even larger incentives to participants is always possible. For example, participants could be offered US \$500 instead of US \$50 if they succeed in convincing the interviewer that they are telling the truth. Introducing severe punishments for those who fail to convince the interviewer that they are telling the truth is not possible, however. University Ethics Committees will not allow researchers to do this. Also, punishments are never realistic, and participants may be aware of it. Ethical guidelines require researchers to inform participants prior to participating that they are free to withdraw from the study at any time. Hence, when participants are threatened with having to enter a dark room to face white noise for one hour, as in Frank and Ekman's (1997) study, they will realise that they are actually free to leave. In other words, it may not be possible to introduce truly high-stakes settings in laboratory experiments, and examining how liars behave in high-stake real-life situations is often the only option (Barrett, 2005; Riggio, 1994).

THE BEHAVIOUR OF A LIAR

The 132 studies published in English examining nonverbal cues to deception in adults that I could find are listed in Appendix 3.1 at the end of this chapter. For each behaviour I have indicated its relationship with deception. A "<" sign indicates that a particular behaviour was shown less by liars than by truth tellers; a ">" sign means that a particular behaviour was shown more by liars than by truth tellers; and a "–" sign means that no difference was found between liars and truth tellers.[8] Empty cells mean that the behaviour was not investigated in that study. In Appendix 3.1 I made a distinction between behaviours that are related to voice and speech (vocal cues, top part) and the other behaviours (visual cues, bottom part). Box 3.2 gives descriptions of these behaviours.

Box 3.2 A description of the nonverbal behaviours

VOCAL CUES

1. *Speech hesitations*: use of speech fillers e.g., "ah", "um", "er", "uh" and "hmmm"
2. *Speech errors*: grammatical errors, word and/or sentence repetition, false starts, sentence change, sentence incompletions, slips of the tongue, etc.
3. *Pitch of voice*: changes in pitch of voice, such as rise in pitch or fall in pitch
4. *Speech rate*: number of spoken words in a certain period of time
5. *Latency period*: period of silence between question and answer
6. *Pause durations*: length of silent periods during speech
7. *Frequency of pauses*: frequency of silent periods during speech

VISUAL CUES

8. *Gaze*: looking into the face of the conversation partner
9. *Smile*: smiling and laughing
10. *Self-adaptors*: scratching the head, wrists, etc.
11. *Illustrators*: hand and arm movements designed to modify and/or supplement what is being said verbally
12. *Hand and finger movements*: movements of hands or fingers without moving the arms
13. *Leg and foot movements*: movements of legs and feet
14. *Trunk movements*: movements of the trunk
15. *Head movements*: head nods and head shakes
16. *Shifting position*: movements made to change seating position
17. *Blinking*: blinking of the eyes

[8] In all Appendices in this book, when I discuss differences between truth tellers and liars, I am referring to differences that are statistically significant.

At the bottom of Appendix 3.1 I also included a summary of the results of DePaulo, Lindsay *et al.*'s (2003) quantitative meta-analysis of cues to deception. The numbers in the bottom cells of Appendix 3.1 represent the effect sizes, d, that were computed by Bella DePaulo and her collaborators. Cohen (1977) suggested that effect sizes of .20, .50, and .80 should be interpreted as small, medium, and large effects, respectively. Negative ds indicate that the behaviour was shown less often by liars than by truth tellers and positive ds indicate that the behaviour was shown more often by liars than by truth tellers. Only the ds printed in bold were significant, that is, indicate a real difference between truth tellers and liars.

The findings reported in Appendix 3.1 show an erratic pattern and indicate that many conflicting results have been found. Take speech hesitations for example. In some studies liars included more hesitations in their speech than truth tellers, whereas in other studies liars hesitated less than truth tellers. In yet another set of studies no relationship between hesitations and deception was found. The same erratic pattern emerged for most of the other behaviours. In other words, a cue akin to Pinocchio's growing nose does not exist. Appendix 3.1 further shows that the patterns for some cues are less erratic than for others. A trend emerges that liars speak with a higher pitched voice than truth tellers. This was also found in DePaulo, Lindsay *et al.*'s meta-analysis, but, in Cohen's terms, a d-value of .21 indicates that the relationship between pitch of voice and deception is weak. Moreover, these differences in pitch between truth tellers and liars are usually very small, only a few Hertz, and therefore only detectable with sophisticated equipment.

Another trend emerges for pauses. Liars seem to include longer pauses in their speech than truth tellers. However, when examining the frequency of pauses, no clear pattern emerged. It thus seems that liars pause longer but no more frequently than truth tellers. DePaulo, Lindsay *et al.* (2003) did not differentiate between frequency and duration of pauses, and the two types of pauses combined were not associated with deception in their meta-analysis ($d = .01$). DePaulo, Lindsay *et al.* (2003) did not find a relationship between latency period and lying. However, Appendix 3.1 shows that in many studies liars waited longer before giving an answer than truth tellers. This was particularly the case in studies where participants took part in reaction time tests and were instructed to respond as quickly as they could. Sporer and Schwandt (2006a), who conducted a meta-analysis about several vocal characteristics, also found that liars show longer latency periods than truth tellers, although the effect size was small, $d = .18$.

In many studies it was further found that liars make fewer illustrators, fewer hand and finger movements, and fewer leg and foot movements than truth tellers. DePaulo, Lindsay *et al.*'s meta-analysis showed that the effect for illustrators was small ($d = -.14$), whereas the effect for hand and finger movements was somewhat more substantial ($d = -.36$). However, most of the effects for hand and finger movements have been found in my own studies, and not many other researchers seem to measure these movements. I would feel more confident about the relationship between hand/finger movements and deception if other researchers could replicate our findings. The effect for leg and foot movements was not significant in DePaulo, Lindsay *et al.*'s meta-analysis, perhaps because a "no difference" was found in many individual studies. However, in those studies where a difference was found, liars typically showed fewer leg and foot movements than truth tellers, as Appendix 3.1 illustrates.[9] Finally, most researchers that obtained an effect for shifting position reported that liars change their sitting position more than truth tellers do. In most studies where shifting position was measured, however, no difference between truth tellers and liars was found. It is therefore probably best to conclude that shifting position is not related to deception. DePaulo, Lindsay *et al.* (2003) did not distinguish between shifting position and trunk movements, but they found that the two cues combined were not related to deception ($d = .05$).

Of the behaviours listed in Appendix 3.1, hand/finger movements appear to have the strongest relationship with deception. Our findings make clear that this cue cannot be treated as Pinocchio's growing nose either. When we analysed our results of 181 participants, we found that 64% of them showed a decrease in hand/finger movements during deception, whereas 36% showed an increase of these movements during deception (Vrij, Winkel, & Akehurst, 1997). Moreover, an effect size of $d = -.36$ should still be considered as small to medium in Cohen's terms.

DePaulo, Lindsay *et al.* included many more behaviours in their meta-analysis than the ones listed in Appendix 3.1. They examined around 100 different nonverbal behaviours. Significant findings emerged for 21 behaviours including three behaviours already discussed: pitch of voice; illustrators; and hand/finger movements. The remaining 18 cues are listed in Table 3.1. Eight of those cues, listed in the bottom half of Table 3.1, were investigated in only a few studies and will not be discussed further. The cues are ranked in terms of their effect sizes.

[9] Also, Sporer and Schwandt (2006b) found in their meta-analysis a significant, but weak, effect ($d = -.13$) for leg and foot movements.

Table 3.1 Nonverbal cues to deception

	d
Pupil dilation	.39
Discrepant/ambivalent	.34
Verbal and vocal uncertainty	.30
Nervousness, tenseness	.27
Vocal tension	.26
Chin raised	.25
Word and phrase repetitions	.21
Verbal and vocal involvement	−.21
Lip pressing	.16
Facial pleasantness	−.12
Cues based on a small number of studies	
Changes in foot movements	1.05
Pupil changes	.90
Genuine smile	−.70
Indifferent, unconcerned	.59
Interrupted words and repeated words	.38
Seeming planned, lacking spontaneity	.35
Intensity of facial expression	−.32
Direct orientation	−.20

Source: derived from DePaulo, Lindsay *et al.* (2003)

The highest effect sizes were found in the cues that have not often been investigated (bottom half of Table 3.1), but if we concentrate on the cues that were investigated more often, then the largest effect size was found for pupil size (dilation) which obtained a *d*-score of .39.

Table 3.1 shows that, compared to truth tellers, liars' pupils look more dilated. In addition, liars appear tenser, have a more tense voice, have their chin more raised, press their lips more, and have less pleasant looking faces. They also sound more ambivalent, less certain and less involved, and make more word and sentence repetitions.

Support for the Theoretical Perspectives

The results provide general support for most of the theoretical perspectives discussed previously. Several cues (pitch, pupil dilation, nervousness, vocal tension, lip pressing, and facial unpleasantness) indicate that liars may be tenser than truth tellers. The findings that liars pause longer, wait longer before giving an answer, make fewer illustrators, fewer hand/finger movements, and fewer leg and foot movements, have larger pupils, and make more word and phrase repetitions suggest that lying is somewhat more cognitively demanding than truth telling. The remaining cues reveal that liars appear more ambivalent, less involved, and more uncertain, and this fits well with the predictions that liars endorse their statements less convincingly than do truth tellers (DePaulo, Lindsay *et al.*, 2003) and that liars often fail to control their behaviour in a convincing manner (DePaulo, Lindsay *et al.*, 2003; Zuckerman *et al.*, 1981a). Rigidity resulting from the decrease in illustrators, hand/finger movements, and leg and foot movements typically exhibited by liars could also be the result of their failing to control their behaviour convincingly.

Further support for Zuckerman *et al.*'s (1981a) multi-factor model, and DePaulo's self-presentational perspective, is obtained via participants' self-reports in both laboratory and diary studies. In some laboratory studies researchers have examined the strategies that truth tellers and liars reported having used during their laboratory interviews. These studies revealed that liars take their credibility less for granted than truth tellers. They therefore try more to make a credible impression, for example by trying harder not to look anxious or nervous (Colwell, Hiscock-Anisman, Memon, Woods, & Michlik, 2006; Granhag & Strömwall, 2002; Granhag *et al.*, 2007; Hartwig, Granhag, & Strömwall, 2007; Strömwall, Granhag, & Landström, 2007; Strömwall, Hartwig, & Granhag, 2006).

In other laboratory studies, after finishing an experimental task, truth tellers and liars are asked to report their experiences during the task. Liars typically indicate that they were more nervous and experienced more cognitive load than truth tellers, and that they tried harder to make a credible impression than truth tellers.[10]

The participants in DePaulo, Kashy *et al.*'s (1996) diary study reported that they felt somewhat more tense when they lied compared to just before telling the lie. We found in our diary study that participants were more nervous when they lied than when they told the truth (Vrij, Ennis, Farman, & Mann, 2006). Diary studies reveal in a different way that people feel somewhat uncomfortable when they lie: self-reports demonstrate that participants tend to lie less in

[10] Caso, Gnisci, Vrij, & Mann, 2005; Gozna & Babooram, 2004; Granhag & Strömwall, 2002; Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Strömwall *et al.*, 2006; Vrij, Edward, & Bull, 2001c; Vrij & Mann, 2006; Vrij, Mann, & Fisher, 2006b; Vrij, Semin, & Bull, 1996; White & Burgoon, 2001. The "nerves" that participants experience in experimental studies are unlikely to be caused by feelings of guilt, because participants are instructed to lie in those experiments and the lies are thus sanctioned by the experimenter.

face-to-face interactions and lie slightly more via a more distant modality (e.g., the telephone) (DePaulo, Kashy *et al.*, 1996; Vrij, Ennis *et al.*, 2006). We further found that when lying, compared to when telling the truth, participants experienced more cognitive load, tried harder to appear credible, and could embrace less what they had said (Vrij, Ennis, *et al.*, 2006).

Despite the support for Zuckerman *et al.*'s and DePaulo's perspectives, it is important to note that those diary studies revealed that the majority of lies were not really taxing. Participants in DePaulo, Kashy *et al.*'s (1996) diary study indicated that their lies were generally not serious, that they put little effort into planning them, and that they did not worry much about the possibility of being caught. Participants further experienced little regret about their lies and said they would tell them again if given a second chance. Our diary study further showed that the participants did not feel much tenseness or cognitive load when they lied, neither did they put much effort in trying to appear credible (Vrij, Ennis, *et al.*, 2006). In summary, most lies told in daily life are low-stakes lies, and it is unlikely that liars will show any clear behavioural signs of emotion, cognitive load, and attempted behavioural control when telling such lies.

Even in some high-stakes situations behavioural signs of deception are unlikely to occur, for example, when a liar gives a false account, that is, an untruthful statement that the liar does not know to be inaccurate. The woman who falsely believes that she was abused in her childhood and who reports her alleged abuse to the police may not show such behavioural signs. She believes she is telling the truth and therefore has no reason to feel guilty. She may also be able to recall her false memory with little cognitive effort.

In contrast, those who believe themselves to be lying, but are in fact inadvertently telling the truth may show behavioural "signs of deceit". In Chapter 2 I gave an example of a suspect who thought that his friend was hiding in his apartment, but, in order to protect his friend, told the police that his friend was abroad. Unbeknown to him, his friend had actually fled the country. Although his statement was therefore factually correct, the suspect himself thought that he was lying to the police. He subsequently may have experienced negative emotions (e.g., guilt or fear) or cognitive load, and may have tried to appear credible while making this statement about his friend.

In the overwhelming majority of deception studies no interaction takes place between sender and target, making them inappropriate to test Buller and Burgoon's (1996) Interpersonal Deception Theory (IDT). Studies in which an interactional interview style has been employed have provided mixed results for the IDT premise that liars adapt their behaviour to diminish suspicions and to appear more credible (DePaulo, Ansfield, & Bell, 1996; Levine & McCornack, 1996a, b). In one study, questioning by the interviewer resulted in an increase in speech disturbances and longer pauses (Buller *et al.*, 1989). However, these behaviours lead to a dishonest demeanour, as will be shown in Chapter 5. In a second study, the interviewer's questioning resulted in an increase in self-adaptors, which does not make an honest impression either (Buller, Strzyzewski, & Comstock, 1991). When in their meta-analysis DePaulo, Lindsay *et al.* (2003) compared deception studies carried out in an interactive context with studies carried out in a non-interactive context, the premise that liars avoid displaying suspicious behaviours could not be supported. It might be that liars aim to suppress all behaviours that they believe arouse suspicion, but they often do not succeed (Buller, Stiff, & Burgoon, 1996). Indeed, as I discussed above, it is often difficult for people to display the behaviours they want to show.

Two Striking Findings

There are two findings that I would like to emphasise, because I believe they often take people, both laypersons and professional lie catchers, by surprise. The first striking finding is that gaze behaviour is not related to deception. This is striking, because across the world laypersons and professional lie catchers believe that liars look away from the target of their deceit. This notion is also frequently mentioned in police manuals, as I will discuss in more detail in Chapter 5 where I will also attempt to explain why people hold the incorrect belief that liars look away. I think there are two reasons why gaze aversion is not a reliable indicator of deception. First, as I mentioned above, gaze has great communicative potential (e.g., eye contact is used to persuade others) and, as a result, people are practised at using and therefore controlling their gaze. Behaviours that are well practised and easy to control are unlikely to be reliable indicators of deception. Second, gaze is related to many factors that have nothing to do with lying. Just to mention a few, people make more eye contact when they interact with people they like than when they interact with people they dislike; they make more eye contact when they interact with people with high status compared to people with low status; they make less eye contact when seated close together rather than at a distance; and also make less eye contact when they feel embarrassed (Kleinke, 1986). Obviously, gaze cannot be reliably related to deception when it is associated with so many factors other than deceit.

Box 3.3 Eye movements and neuro-linguistic programming

Many police officers believe that specific eye movements provide information about whether a person is lying (Leo, 1996a) and this idea is sometimes mentioned in the police literature. In particular the belief that liars look up and to the left seems to be popular. This idea is derived from the neuro-linguistic programming model, although those who developed this model never referred to a relationship between such eye movements and deception. Evidence that eye movements indicate deception is lacking. Even those who suggest that this relationship exists have never presented data supporting their view. See Vrij and Lochun (1997) for a critical review of this police literature.

The second striking finding is that experiencing emotions and cognitive effort, or attempting to appear credible, is not exclusive to liars. It may also happen to truth tellers. As a result, both truth tellers and liars may show signs of emotions, cognitive effort, or behavioural control. Suppose that, when asked whether he killed his wife, a suspect answers "No, I did not do that" with a high-pitched voice that sounded tense. Is this man lying? That is impossible to say. Apparently he shows nervous behaviour, but it is unclear why he is nervous. He may be nervous because he is trying to hide the fact that he killed his wife. However, he also may be nervous *when he is telling the truth*, for instance, because he is afraid that the police detective will not believe him when he says that he did not kill his wife.

Suppose that a suspect stops fidgeting as soon as she starts mentioning an alibi. This may indicate that she is experiencing cognitive load, but it is impossible to say whether she is lying. She may have stopped fidgeting because she is lying and is having real difficulty in fabricating a convincing alibi. However, she may also stop fidgeting if innocent. She may realise that her alibi is an important element in her defence, and therefore thinks hard to be sure that she tells every detail correctly and does not leave out information that may be important.

I believe that police officers may not be sufficiently aware of the fact that truth tellers and liars may experience similar emotions and cognitive processes. In discussions with police detectives, they often come up with statements such as: "I am sure that he is lying, because he appeared nervous when we discussed the crime." It is often too premature to draw this conclusion on the basis of the suspect's behaviour. In most cases, a lie detector has to undertake further action, such as asking more questions or checking the information the interviewee provides, in order to find out whether the person is lying. Drawing conclusions about deception solely on the basis of someone's behaviour is often not reliable. I will return to this issue later in this book.

REASONS FOR FEW NONVERBAL CUES TO DECEPTION

The theoretical perspectives, outlined above, already predicted that research would reveal only a few, and usually weak, relationships between nonverbal cues and deception. There are more explanations for the limited success that researchers have in pinpointing nonverbal cues to deception. This section highlights several of these reasons.

Inadequate Scoring Systems

One explanation for not finding consistent and reliable nonverbal cues to deception is that the scoring systems used to measure them are not detailed enough. We already saw two examples of this. The category "frequency and duration of pauses combined" was not related to deception, but, when examining "frequency" and "duration" separately, it appeared that liars pause somewhat longer but no more often than truth tellers. Also, the total category of speech errors that includes grammatical errors, stutters, false starts, slips of the tongue, and so on did not reveal a relationship with deception, but a subcategory of speech errors – repetition of words or phrases – was related to deceit with liars repeating themselves more than truth tellers. The same relationship between word repetitions and deception was found by Davis, Markus, Walters, Vorus, and Connors (2005) in their analysis of suspects' confessions, and also in the interview with Ian Huntley, the British caretaker convicted for murdering two 10-year-old girls. I will discuss those cases later on in this section.

Although not yet investigated, a similar pattern may emerge if we examine in more detail another vocal category, speech hesitations ("ah", "um", "er", "uh", and "hmmm"). Smith and Clark (1993) found that "um" indicates higher cognitive load than "uh". Therefore examining "ums" and "uhs" separately may be indicative of lying. Perhaps liars include more "ums" in their statements than truth tellers, due to enhanced cognitive load, but do not differ in their use of "uhs".

It is perhaps not just vocal characteristics that could have suffered from inadequate scoring systems. Ekman (1985/2001) has identified a number of different smiles, including a distinction between felt and false smiles. This distinction is useful to detect deception. Felt smiles include smiles in which the person actually experiences a positive emotion. Felt smiles are accompanied by the action of two muscles: the *zygomatic major* that pulls the lip corners upward towards the cheekbone and the *orbicularis oculi* that raises the cheek and gathers skin inwards from around the eye socket. The latter change produces bagged skin below the eyes and crow's-feet creases beyond the eye corners. False smiles are deliberately contrived to convince another person that a positive emotion is felt, when in fact it is not. In false smiles, the action of the orbicularis oculi muscle causing the effects around the eye is often missing (Frank, Ekman, & Friesen, 1993). Ekman and colleagues found that truth tellers make more felt smiles than liars, whereas liars make more false smiles than truth tellers. When the distinction between felt and false smiles is not made, truth tellers seem to smile as frequently as liars (Ekman, Friesen, & O'Sullivan, 1988). Other differences between felt and false smiles include that false smiles are more asymmetrical, appear too early or too late, often last longer, and have a less consistent duration (Ekman, 1988; Ekman, Davidson, & Friesen, 1990; Ekman & Friesen, 1982; Ekman & O'Sullivan, 2006; Frank, Ekman, & Friesen, 1993).[11]

Researchers perhaps also failed to notice some specific movements liars make. Nonverbal communication researchers have identified numerous types of hand movements (Efron, 1941; Ekman & Friesen, 1969, 1972; Rime & Schiaratura, 1991). Based on the work of Efron (1941), Ekman and Friesen (1969) made a distinction between five hand movement categories: emblems, illustrators, affect displays, regulators, and self-adaptors. In their later writings (Ekman & Friesen, 1972; Friesen, Ekman, & Wallbott, 1979), they restricted themselves to only three of these categories, emblems (gestures with a specific meaning like "thumbs up"), illustrators, and self-adaptors, because "these three classes include all hand movements except for those times when the hand moves simply to establish a new position or rest" (Friesen *et al.*, 1979, p. 99). This three-class categorisation is used in deception research, and all three categories appeared in DePaulo, Lindsay *et al.*'s (2003) meta-analysis. (Emblems were not a diagnostic cue to deceit in DePaulo's meta-analysis, but very few researchers have investigated them.)

[11] A smile with the action of the orbicularis oculi muscle is not always felt. Actions of this muscle may occur during false smiles, as the muscle is also involved in other felt emotions, such as distress, sadness or pain. When somebody tries to hide those feelings with a false smile, actions of the orbicularis oculi may occur. Moreover, good facial performers may successfully manage to produce false smiles that look like felt smiles.

Ekman and Friesen (1972) make further distinctions into eight types of illustrators, but these subdivisions are typically not used by deception researchers. In one experiment, however, we did differentiate between different types of illustrators (Caso, Maricchiolo, Bonaiuto, Vrij, & Mann, 2006). Truth tellers described objects they had in their possession, whereas liars had to imagine that they had these objects in their possession. Liars made fewer deictic movements (pointing gestures) than did truth tellers, perhaps due to the lack of real objects they could point at, but liars made more metaphoric gestures that are typically made when people describe abstract ideas (McNeill, 1992). When we examined all illustrators combined, no relationship with deception emerged. In other words, like word repetitions and smiles, it was only when specific types of distinctions were made among subclasses of illustrators that deception cues emerged.

Finally, it may be that researchers do not define the truthful and deceptive fragments they code in enough detail to reveal cues to deceit. For example, eye blinks decrease as a result of cognitive load, but during gaps of cognitive demand a flurry of blinks occurs (Holland & Tarlow, 1972; Leal, 2005; Stern, Walrath, & Goldstein, 1984). This means that how the deceptive fragments in lies associated with high cognitive load are selected is an important consideration. It may well be that such lies are associated with a decrease in eye blinks but that a flurry of blinks occurs *directly after* they are told. Researchers who include the period directly after the lie in their deceptive fragment may therefore overlook the decrease in blinking that occurred during the actual lie (Fukuda, 2001).

In summary, all these findings converge to the idea that more diagnostic cues to deception could be found if observers examine nonverbal responses in more detail. Observers should avoid lumping together different types of behaviour (such as different types of speech errors, hesitations, smiles, and illustrators) and should precisely define what constitutes a deceptive fragment. It may further make a difference whether they measure the frequency of occurrence or the duration of each behaviour (e.g., pauses).

Box 3.4 Micro-expressions of emotions

Ekman believes that observing emotional micro-expressions in the face may reveal valuable information about deception (Ekman, 1985/2001). Emotions that are strongly felt almost automatically activate muscle actions in the face. For example, anger results in a narrowing of the lips and lowering of the eyebrows; eyebrows that are raised and pulled together and a raised upper eyelid and tensed lower eyelid typically mark fear; and joy, as mentioned in the main text, activates muscles that pull the lip corners up, bag the skin below the eyes, and produce crow's-feet wrinkles beyond the eye corners. If a person denies an emotional state that is actually being felt, he or she will have to suppress these facial expressions. Thus if a scared person claims not to be afraid, that individual has to suppress the facial micro-expressions that typically indicate fear. This is difficult, especially because these emotions can arise unexpectedly. For instance, people do not usually deliberately choose to become frightened, this happens automatically as the result of a certain event that took place, or as the result of a particular thought. Suppose that a suspect who claims to be relaxed becomes frightened during a police interview at the moment when she finds out that the police know more about her involvement in the crime than she thought they would know. The moment fright occurs she may show a fearful facial expression. There is a chance that she will try to suppress this emotional expression in order not give away that she is afraid. People are usually able to suppress these expressions quickly, within 1/25th of a second after they begin to appear (Ekman, 1985/2001). Nevertheless the expression is present for a short period of time and can be spotted by trained observers (Ekman, 1985/2001).[12]

Rather than attempting to mask a felt emotion, someone can try to do the opposite: pretending to experience a particular emotion, when in fact this emotion is not felt. A mother can pretend to be angry with her child when in reality she is not angry. In order to be convincing, the mother should produce an angry facial expression and should try to narrow her lips. This muscle action is very difficult for most people to make voluntarily (Ekman, 1985/2001).

Finally, someone can pretend an emotion other than the one that is actually felt. An adulterous husband may become scared during a conversation with his wife when he realises how much she knows about his affair, but can decide to mask this emotional state by pretending to be angry with his wife because she apparently does not trust him. In order to be convincing, he therefore has to suppress his fearful facial expression and replace it with an angry facial expression. This is difficult, because he has to lower his eyebrows (sign of anger) whereas his eyebrows tend to rise (sign of fear) (Ekman, 1985/2001).

[12]In this respect the recent development of computer-based automated analyses of facial expressions may be helpful (see Masip, 2006).

Similar to the findings with smiles, spontaneous expressions of emotions may differ from deliberate expressions in ways other than the activation of different muscles. For example, it has been found that spontaneous and deliberate expressions differ in latency time, onset time (time from the start of the expression to the peak of intensity of that expression), offset time (time from the first evidence of fading of the intensity of the expression until the disappearance of the expression), duration of peak intensity, and overall duration (Ekman, Friesen, & Simons, 1985; Hess & Kleck, 1990; Hill & Craig, 2002; see also Ekman & O'Sullivan, 2006).

A Combination of Cues

It could be that when researchers examine each nonverbal cue individually, no diagnostic cue to deception occurs, but that a diagnostic pattern does arise when a combination of cues is taken into account. There is some evidence for this. For example, in our own experiment we found that on the basis of a combination of four nonverbal behaviours (illustrators, hesitations, latency period, and hand/finger movements) we could correctly classify 70.6% of participating truth tellers and 84.6% of liars, whereas any of these behaviours individually resulted in much more disappointing results (Vrij, Edward, Roberts, & Bull, 2000).

Ekman and colleagues found a similar pattern. Up to 80% of truths and lies could be detected when a trained observer paid attention to micro-facial expressions (Frank & Ekman, 1997), but even better classifications of truth tellers and liars were obtained when in addition to micro-facial expressions the tone of voice was taken into account. In that situation 86% of the truths and lies could be detected (Ekman, O'Sullivan, Friesen, & Scherer, 1991). In other studies between 71% and 78% of correct classifications were made when the researchers investigated a cluster of behaviours (Davis *et al.*, 2005; Heilveil & Muehleman, 1981; Vrij, Akehurst *et al.*, 2004a). In other words, clusters of behaviours may show diagnostic patterns to deceit. However, this raises the question of *which* behaviours should be clustered. At present, different researchers examine different clusters of behaviour, and it cannot be ruled out that a cluster that is effective in pinpointing lying in one situation or group of participants is not effective in another situation or group of participants. However, as a principle, more accurate truth/lie classifications can be made if a cluster of nonverbal cues is examined rather than each of these cues individually.

Perhaps this idea of investigating a cluster of cues rather than individual cues could also explain why concepts such as ambivalence, uncertainty, involvement, nervousness, tension, and pleasantness all emerged as cues to deceit (see Table 3.1). Making assessments of such states is likely to be based on a cluster of behaviours rather than on individual cues.

Group Differences in Cues to Deception

It could be that different people show different nonverbal cues to deceit. In this section I discuss these so-called *interpersonal* differences with regard to an individual's ethnicity or culture, gender, age, and personality traits.

Ethnic origin or culture

It could be that people of different ethnicities or cultures show different nonverbal cues to deceit. For example, perhaps Caucasian participants display different give-away cues than participants from another ethnic origin or culture. Unfortunately, there is not much cross-cultural deception research. Most deception studies have been carried out in Western countries where the vast majority of participants were of Caucasian origin. In other words, the research findings published to date tell us how Caucasian liars behave. Although the lack of cross-cultural research is a shortcoming, one should not become too excited that different results will arise with non-Caucasian participants. It is important to keep in mind *why* nonverbal cues occur. They occur because the liar experiences emotions or cognitive load, or tries to behave in a convincing manner. There is no theoretical reason why those experiences will be different across cultures or why manifestations of emotions, cognitive effort, and attempted behavioural control will differ between cultures. Not surprisingly, in the few studies where cues to deception in participants of different ethnic groups were compared, no differences between ethnic groups were found (Sitton & Griffin, 1981; Vrij & Winkel, 1991).[13]

[13] Showing the same cues to deceit is not the same as showing the same patterns of behaviour. People from different ethnic backgrounds or cultures often display different behavioural patterns (Matsumoto, 2006). For example, Caucasians tend to look conversation partners in the eye more than African Americans (Fugita, Wexley, & Hillery, 1974; Ickes, 1984; Johnson, 2006; LaFrance & Mayo, 1976; Smith, 1983). However, truth tellers and liars do not differ in eye contact in Caucasian participants, neither do truth tellers and liars differ in eye contact in those non-Caucasian groups that have been examined (Sitton & Griffin, 1981; Vrij & Winkel, 1991). In other words, gaze behaviour is influenced by culture, but not by veracity. Differences in behaviour displayed by people from different ethnic origins or cultures may easily lead to errors in lie detection as I will discuss in Chapter 6.

Gender

Gender differences between truth tellers and liars are unlikely to occur for the same reason that ethnic or cultural differences are unlikely to occur. In many situations men and women do not differ in emotions felt when lying, in the cognitive load they experience when they lie, or in their attempts to give a convincing impression to observers; nor are there any theoretical reasons why the behavioural manifestations of emotions, cognitive effort, and attempted behavioural control would differ between males and females. Perhaps for this reason, researchers rarely report gender differences in their deception research, although in the vast majority of studies both men and women have participated.[14]

Age

A different picture may emerge when considering the age of participants. There are theoretical arguments why children may reveal different cues to deceit than adults, although this theoretical picture is not straightforward. The attempted behavioural control approach suggests that liars, first, should realise that observers will watch their responses to detect deceit, second, should know which responses make an honest impression on others, and, third, should be able to display the desired responses. The first two aspects imply that the effective liar should be able to "take the role of the other". This skill is largely lacking in children under six years old (Broomfield, Robinson, & Robinson, 2002; Flavell, 2000; Flavell, Botkin, Fry, Wright, & Jarvis, 1968) and probably contributes to adults' views that children are less socially aware and less adept at communication than adults (Leippe, Brigham, Cousins, & Romanczyk, 1987; Melinder, Goodman, Eilertsen, & Magnussen, 2004). Others refer in this context to the Theory of Mind (Gallup, 1982; Keenan, Gallup, & Falk, 2003; Sabbagh, Moses, & Shiverick, 2006), which is the capacity to understand that other people's thoughts differ from one's own thoughts, and that others do not necessarily know what thoughts are going through one's own mind (Johnson, Barnacz *et al.*, 2005). Younger children's lesser understanding of someone else's mental states suggests that they will not try so hard to suppress nonverbal cues that may betray their lie. As a result, more cues of nervousness and cognitive demand could be expected in younger children when lying than in adults.

[14] See Hall (1984, 2006) for reviews of gender differences in behaviour.

Box 3.5 Theory of mind and skills in lying

Research has demonstrated that as they get older, children become better at influencing other people's mental states. In a hiding game in which children had to deceive a competitor, Sodian (1991) found that three year olds used acts of "sabotage" (physically preventing the competitor from gaining a reward) whereas most four year olds fooled the other person by pointing in the wrong direction to prevent that person from stealing the object. In another hiding game, in which children had to point to an empty box in order to deceive another, Russell, Mauthner, Sharpe, and Tidswell (1991) found that four year olds performed remarkably better than three year olds. In Peskin's (1992) study, three to five year olds were confronted with a competitor who always chose the object for which the children themselves had previously stated a preference. The children were asked to "think of what to do or to say so that the competitor won't get what you want". In the game, the competitor asked each child: "What do you want?" More five year olds (87%) than three year olds (29%) tried to influence the competitor's mental state by pointing to an object they did not like or by concealing information ("I won't tell you").

Children's muscular control also increases with age (Ekman, Roper, & Hager, 1980; Feldman & Phillipot, 1993; Hala & Russell, 2001; Kieras, Tobin, Braziano, & Rothbart, 2005, Saarni, 1984; Sabbagh, Moses, & Shiverick, 2006; Talwar, Murphy, & Lee, 2007). For example, in a naturalistic study, Saarni (1984) observed 6–10 year olds' reactions when they received a disappointing gift. Younger children were more likely than older children to show their true reaction (negative faces). Ekman *et al.* (1980) studied 5, 9, and 13 year olds and found that older children have a greater ability to deliberately produce the component actions involved in facial expression. These findings suggest that as age increases, cues to deceit are likely to decrease, because children are able to control themselves better as they get older.

The arguments presented so far would suggest that young children are poorer liars than adults. However, three more arguments cloud this picture. DePaulo and Jordan (1982) have argued that younger children may experience less emotion when lying. For example, because of their younger years they may be more inclined to overlook the negative consequences of being detected and so experience less fear of getting caught. Also, with increasing age, people show more spontaneous emotional facial expressions, which they sometimes will need to suppress in order to conceal deceit (Morency & Krauss, 1982). Therefore, adults' role-play skills and increased muscular control, which theoretically should make them more skilful liars, may well be counteracted in part by an increase in experiencing emotions while lying and by an increase in spontaneous emotional expression. Second, as already discussed above, truth tellers also may experience emotions and cognitive demand, which suggests that at the very best only small differences will emerge between young truth tellers and liars in terms of emotion and cognitive demand. Third, young children typically provide shorter statements than adults when they are invited to recall an event they have experienced (Lamb, Orbach, Sternberg, Esplin, & Hershkowitz, 2002). This may benefit young children when they lie, because when liars talk for a shorter period of time, there is less opportunity for them to display nonverbal cues to deceit, and the less likely it is that they will show such cues (DePaulo, Lindsay *et al.*, 2003).

Studies examining how children behave when they lie are scarce. This is partly due to ethical constraints. In experimental studies participants are often instructed to lie. This is problematic with children because they are typically taught at school and at home not to lie. Researchers get around this issue by observing children's spontaneous lies or by instructing children to tell white lies. In the studies examining behavioural cues to deception published to date the stakes were low, the lies were easy to tell, and there was no effort made to motivate the children to lie (Feldman, Devin-Sheehan, & Allen, 1978; Lewis, 1993; Lewis, Stanger, & Sullivan, 1989; Talwar & Lee, 2002; Vrij & Winkel, 1995). For example, in both Lewis' studies, children (two and three year olds) were instructed that they must not peek at an object that was behind them. However, some of the children did peek and denied having done so (thus lied) when asked. Their responses were compared with those children who did not peek and therefore truthfully responded that they did not peek. In the study conducted by Feldman and colleagues, eight year olds were asked to praise a confederate who conducted a task. In the truthful condition the confederate performed well and the praise was therefore honest. In the deception condition the confederate performed poorly and the praise was therefore deceptive. All of these studies only revealed small differences between truth tellers and liars, and a comparison between different studies did not show a consistent pattern regarding which cues revealed deceit (Vrij, 2002a). Interestingly, in none of these studies were gaze aversion or fidgeting, the cues people typically associate with deception (Chapter 5), actually related to deception.

The studies mentioned so far in this section only examined the responses of children. On the basis of such studies, we cannot really know whether children show different nonverbal cues to deception than adults. To examine this, children's and adults' reactions in the same experiment need to be compared. To my knowledge we are the only researchers who directly compared adults' (university students) and children's (5–6 year olds) nonverbal responses while lying (Vrij, Akehurst *et al.*, 2004a). Participants in the truthful condition played a game of Connect 4[15] with a confederate. During the game, someone came in, made a comment, wiped the white board, and walked out again. Liars did not actually participate in this event but were informed about what happened. They were then instructed to pretend that they had experienced the event in a subsequent interview. Both truth tellers and liars were interviewed about the event. Participants were promised a small incentive if they could give a convincing impression during the interview. The interviewer did not know who was telling the truth and who was lying. Large differences were found in the behavioural responses of adults and children. For example, the children made almost twice as many movements as the adults. However, deceptive cues were remarkably similar. In both adults and children, deception was associated with a decrease in hand and finger movements. The other cues that we examined, gaze aversion, latency time, pauses, speech rate, speech hesitations, speech errors, self-adaptors, illustrators, and leg and foot movements, did not differentiate between truth tellers and liars in either children or adults.

In summary, not a single study to date has revealed large differences in nonverbal behaviour in children when they tell the truth compared to when they lie. In that respect, the findings resemble those of adults. In one study where the behaviours of adults and children were compared, the same cue to deceit (decrease in hand and finger movements) emerged in both age groups. Popular stereotypical cues to deception, such as gaze aversion and fidgeting, have not been found to be diagnostic cues to deceit in children.

Personality traits

The nonverbal cues to deception people display may be influenced by their personality. However, the findings are not always as clear-cut as

[15] Connect 4 is a two-player game where players drop counters into a slotted grid to achieve, and simultaneously prevent their opponent from achieving, four of their counters in a row.

you may expect. As already discussed, Machiavellians consider lying to be a legitimate way of achieving their goals (Chapter 2), and therefore Machiavellians feel less guilty when they lie (this chapter). They are also more likely to engage in strategic self-presentation to influence others (this chapter). You may therefore expect that people high in Machiavellianism display fewer cues to deception than people low in Machiavellianism, but there is hardly any evidence to support this assumption. Exline, Thibaut, Hickey, Gumpert (1970) did find that people high in Machiavellianism kept more eye contact when lying than those low in Machiavellianism (which may have been caused by feeling less guilty or by more strategic self-presentation), but other studies did not reveal behavioural differences between people high and low in Machiavellianism when lying (Knapp, Hart, & Dennis, 1974; O'Hair, Cody, & McLaughlin, 1981).

In another study, the nonverbal behaviours of people scoring high or low in psychopathy were compared, but no clear differences emerged between the two groups (Klaver, Lee, & Hart, 2007; see also Book, Holden, Starzyk, Wasylkiw, & Edwards, 2006, and MacNeil & Holden, 2006).

As discussed above, people who score high in Public Self-Consciousness particularly see themselves as the focus of other people's attention and try particularly hard to control their behaviour so that they make a favourable impression. Research has revealed mixed findings about whether people high in Public Self-Consciousness actually have more control over their nonverbal presentation style than those low in Public Self-Consciousness. Baumeister (1984) and Gallaher (1992) reported that they have, but we could not replicate this finding in a deception experiment (Vrij, Edward, & Bull, 2001c). In our study we measured to what extent liars displayed suspicious behaviour, such as a long latency period, gaze aversion, and more self-adaptors. We found no relationship between Public Self-Consciousness and displaying such cues. In their deception study, Riggio, Tucker, and Widaman (1987) found that people high in Public Self-Consciousness were poor liars, for example, by exhibiting less eye contact and more emotional reactions than those low in Public Self-Consciousness. In other words, people's intention to avoid displaying suspicious behaviour does not automatically mean that they will succeed in doing this. We saw this earlier when I reviewed the support for Buller and Burgoon's Interpersonal Deception Theory.

There is stronger empirical support that extraverts and introverts show different behaviour when they lie: extraverts show different and fewer cues to deception (Miller, deTurck, & Kalbfleisch, 1983; Riggio & Friedman, 1983; Siegman & Reynolds, 1983). Extraverts display fewer movements when they lie than when they are honest, whereas introverts make more movements when they lie than when they are honest. Introverts also have more disturbances in their speech during deception than extraverts. Introverts usually feel more uncomfortable in social interactions than extraverts, and perhaps lying makes them more nervous than extraverts, resulting in an increase in movements and nonfluencies in speech. Siegman and Reynolds (1983) found that introverts showed more differences in pauses, speech rate, and latency time between truth telling and lying than extraverts, again, perhaps because introverts feel more uncomfortable in social interactions than extraverts.

The ability to act also influences someone's behaviour while lying. Compared to bad actors, good actors show more often a decrease in movements when they lie, perhaps because they may be better at suppressing signs of nervousness (Vrij, Akehurst, & Morris, 1997). Compared to bad actors, good actors also include fewer pauses in their speech when they lie (Miller, deTurck, & Kalbfleisch, 1983; Siegman & Reynolds, 1983). Finally, in a study where participants were put in a situation where it was inappropriate to show happiness, participants high in self-monitoring (a trait that includes both extraversion and acting skills) were better at hiding their happiness than participants low in self-monitoring (Friedman & Miller-Herringer, 1991).

Idiosyncratic Patterns

The previous section about group differences in cues to deception revealed some patterns, particularly regarding introverts/extraverts and good/poor actors. Other classifications, however, still do not result in clear-cut patterns of nonverbal cues to deceit. Perhaps such categorisations are still too broad to reveal such cues, and perhaps only cues to deception emerge when we examine them at an individual level. That is, perhaps such cues are idiosyncratic, and perhaps each individual has his or her own unique set of nonverbal cues which betray his or her lie. Indeed, our analysis of the behaviour displayed by a murderer during his police interview (discussed in detail below) revealed that the man displayed cues of deceit (Vrij & Mann, 2001a), and it is unlikely that another individual would have displayed exactly the same pattern of behaviour in this situation. To further complicate the issue, cues are also situation specific, and there is no guarantee that the murderer would have shown the same cues to deceit if he had been interviewed at a different time, or about a different topic, or outside a police interview setting. In other words, situational factors may also impact upon a person's behaviour, and I will discuss such factors next.

Situational Factors

Other than differences between individuals (which I labelled *interpersonal* differences), differences in the context wherein the lie is told may have an effect on cues to deception (which I labelled *intrapersonal* differences). I will introduce four contextual influences: complexity of the lie, motivation of the liar to get away with the lie, the stakes, and the interview style to which the liar is exposed.

Content complexity of the lie

Sometimes it is difficult to tell a lie. Suppose that an applicant did something very stupid in her previous job and that, to her own surprise, a member of the selection board refers to this stupidity during the selection interview. The applicant, who really wants the job, does not want to admit this blunder and must therefore instantly come up with a fabricated but plausible explanation. This will not be easy. She will probably have to think hard which may well result in an increase in latency period, speech hesitations and speech errors, a slower speech rate, and a decrease in movements.

Liars, however, are not always taken by surprise. They often know what kind of questions they can expect and can therefore prepare themselves. Many guilty suspects, for example, will be aware of the possibility that the police will interview them one day about their activities at the time the crime was committed. They therefore could prepare an answer to this question well in advance. Lying may not be too difficult in this situation. When the police ask him about his activities at the time of the crime, the well-prepared guilty suspect simply gives the answer that he had prepared previously.

How do liars behave when they have had the opportunity to plan their lies? There is evidence to suggest that spontaneous lies are preceded by longer latency periods than spontaneous truths, but that planned lies are preceded by shorter latency periods than planned truths. Moreover, lies in spontaneous messages are associated with more pauses than lies in planned messages (DePaulo, Lindsay *et al.*, 2003). In summary, it is more difficult to tell spontaneous lies than planned lies, resulting in more vocal signs of thinking hard in spontaneous lies.[16]

Sometimes a liar does not have to fabricate an answer at all, but simply has to conceal some information. When a customs officer asks someone what is in the suitcase, the only thing a smuggler has to do

[16] The cue that emerged in planned lies (a shorter latency period) could also be the result of attempted behavioural control. Perhaps liars think that waiting too long before giving an answer will look suspicious and hence a quick response is likely to occur.

is to conceal some information, that is, not mention the contraband. Some of my own studies examined how liars behave in such a situation (Akehurst & Vrij, 1999; Vrij, 1995; Vrij, Akehurst, & Morris, 997; Vrij, Semin, & Bull, 1996; Vrij & Winkel, 1991). In these studies, participants in the deception condition had to deny the possession of a set of headphones they actually possessed. In other words, their task was to conceal some information. Their answers were compared with the answers given by participants who were not in possession of a set of headphones and therefore were honest when they denied the possession of these headphones. In these studies, the liars made fewer speech hesitations and spoke faster than the truth tellers. The following explanation sounds plausible. Not mentioning the set of headphones is perhaps an easy lie to tell, not resulting in cues of cognitive load. Instead, in an attempt to make an honest impression, liars may have therefore tried to avoid making speech hesitations or speaking too slowly, which resulted in the unusually smooth speech pattern and unusually fast speech rate.

In another study, we further examined the moderating impact of lie complexity on the occurrence of speech hesitations (Vrij & Heaven, 1999). In line with the attempted behavioural control approach we hypothesised that liars would try to avoid making speech hesitations. We expected that they would only achieve this when the lie is easy, and that when lying is difficult signs of cognitive load would occur. Participants were shown a video of a couple having an argument. First, the man appears maintaining that he wants to buy satellite television so that he can watch football at home rather than having to go to the pub to watch it. The woman follows explaining that the only reason he really wants to buy satellite television is so that he can bring his friends home from the pub to watch the porn channels. After watching this videotape, the participants were requested to lie about some aspects of the video and tell the truth about other aspects of the video. One lie, giving an inaccurate description of the appearance of one of the people in the video, was relatively easy to fabricate, whereas the other lie, making up a reason why the person in the video wanted to buy satellite television, was more difficult to fabricate. As expected, liars made more speech hesitations (compared to truth tellers) when the lie was cognitively difficult and made fewer speech hesitations (compared to truth tellers) when the lie was easy.

The motivated liar

Liars are not always equally motivated. A murderer in a police interview is probably more motivated to hide the truth than the woman who tells her friend that she likes her friend's new dress. People who are highly motivated to get away with their lies may behave differently from people who care less about the outcome. In their meta-analysis, DePaulo, Lindsay *et al.* (2003) found that studies where some incentive for succeeding was provided (such as money) resulted in more nonverbal cues to deceit than studies where no incentive was provided. Moreover, nonverbal cues to deception were more prominent in studies where participants lied to cover up bad behaviours, such as cheating and stealing (e.g., transgressions) than in studies where they lied about issues such as their age or their opinions (e.g., non-transgressions). In short, it appears that the more motivated liars are to avoid getting caught, the more likely it is that their behaviour will give their lies away. DePaulo and Kirkendol (1989) labelled this phenomenon the *motivational impairment effect*.

This effect may sound surprising, but is easy to explain. Highly motivated liars probably experience stronger emotions than unmotivated liars, for example more fear of getting caught. Additionally, motivated liars probably experience more cognitive load than unmotivated liars. They may put more effort into telling their stories, and may scrutinise the behaviour of the lie-target more carefully to find out whether they get away with their lie. As I discussed above, the stronger these processes are present in liars, the more likely it is that cues to deception occur. Motivated liars probably also try harder to make a credible impression. However, this does not necessarily benefit their performance, because people are often poor at controlling all facets of their behaviour. Often it is the case that the harder they try to control all facets, the worse their performance becomes (Baumeister, 1984; Baumeister & Showers, 1986).

High-stakes lies

Sometimes the stakes are high. In a murder inquiry it really matters to suspects who deny any involvement in the killing, that they are believed. What nonverbal cues to deception occur in high-stakes situations? As I already discussed, laboratory studies cannot provide the answer because the stakes in such studies are never as high as the stakes are for smugglers, guilty suspects, adulterous spouses, fraudulent businessmen, and corrupt politicians when they are interviewed about their activities. To examine how truth tellers and liars respond in high-stakes situations, one of the few options available to researchers is to analyse real-life high-stakes settings.

Few real-life observations have been published. I will describe the analyses I am aware of, and this includes observations of Bill Clinton, Saddam Hussein, and Ian Huntley (a British caretaker found guilty of murdering two 10-year-old girls). Unfortunately, some of the analyses gain little insight into deception because they suffer from problems with establishing the ground truth (Bill Clinton)[17] or selecting comparable truths (the Saddam Hussein and Ian Huntley cases in particular). Other analyses, however, are more valuable. I believe a trend emerges from these real-life observations. It appears that when people lie in high-stakes situations, cues of cognitive load in particular, such as long pauses, word repetitions, or a decrease in hand/finger movements are most likely to occur. This is in direct contrast with the view of professional lie catchers who overwhelmingly believe that liars in high-stakes situations will display cues to nervousness, particularly gaze aversion and self-adaptors (Chapter 5). The analyses presented below show no evidence for the occurrence of such cues.

Bill Clinton When former US President Bill Clinton testified before the Grand Jury on 17 August 1998 about his alleged affair with Monica Lewinsky, he had to answer questions about Betty Currie (his personal secretary) on several occasions. Betty Currie went to Monica Lewinsky's home to collect the presents she had received from President Clinton. The question is whether or not President Clinton instructed her to do this. This is an important question, as it would be a clear sign of an obstruction of justice if Clinton indeed gave instructions. Prosecutor Kenneth Starr's team asked Clinton twice whether he gave Betty Currie such instructions. Clinton denied doing so both times, but each time he showed remarkable behaviour (Vrij, 2002b). Both times he sat up, did not move, and looked straight into the camera. The first time especially, his behaviour was striking. He denied quickly, even before the interviewer had completed his question, and continued his rigid behaviour and looking into the camera during the period of silence that followed after his denial. To me, it appears that he really wanted to make an honest impression on Kenneth Starr's team and the Grand Jury during this particular part of the interview. Unfortunately, I don't know whether he was lying about instructing Betty Currie, because I don't know the ground truth.

[17] Back in 1953, Reid and Arther published a study regarding the behaviour of 486 guilty and 323 innocent people who were suspected of various criminal offences. Their analysis revealed several indicators of deception. However, no details were given about the ground truth in these cases. Horvath (1973) included 100 suspects in his study of which 50 were truthful and 50 were lying. Suspects' speech content and nonverbal communication were observed during a pre-test interview stage of polygraph examinations. Analyses of their nonverbal behaviour revealed several diagnostic cues. Again, no information about the ground truth has been provided. Without certainty about the ground truth, we cannot actually tell whether the observed differences were caused by the fact of lying. I therefore prefer not to discuss these studies in further detail.

Box 3.6 Clinton's Grand Jury testimony

Hirsch and Wolf (2001) do not appear to worry about the ground truth in Clinton's Grand Jury testimony. They observed 23 nonverbal and verbal cues displayed by Clinton during his testimonial. They examined a 23-minute segment of the videotape and compared this with 11 minutes of the same testimony when Clinton answered basic questions (his name, the attorney's name, etc.). Significant differences were obtained for 19 cues. They also compared the 23-minute segment with five minutes of a fundraising speech to a sympathetic crowd. This time, 20 significant differences emerged. Unfortunately, this study tells us nothing about cues to deception. First, the ground truth in the 23-minute segment has not been established. Even when we assume that Clinton was lying in this 23-minute fragment, it is unlikely that he lied continuously during those 23 minutes, because longer deceptive messages typically contain some truthful elements (Turner, Edgley, & Olmstead, 1975; Maclin & Maclin, 2004; Wang, Chen, & Atabakhsh, 2004). Second, unfair comparisons between truthful and deceptive statements were made. Comparisons between this 23-minute fragment and the other fragments (answering basic questions during the interview, or fundraising speech) are apple–orange comparisons. People show different behaviours during small talk at the beginning of the interview than during potentially embarrassing questions about an affair. It is also obvious that people show different behaviours when they address a crowd in a fundraising speech than when they are interviewed about an alleged affair. In this respect, it might be more surprising that significant differences were found for only 19 or 20 cues and not for all 23 cues.

Saddam Hussein Iraq's former President Saddam Hussein was interviewed by the journalist Peter Arnett, at that time working for Cable News Network (CNN), during the first Gulf War (1991). The interview lasted 94 minutes and was broadcast on CNN. Davis and Hadiks (1995) observed and scored Hussein's behaviour during this interview. They used a detailed scoring system to measure every single movement of his hands, arms, and trunk. The scoring method was much more detailed than those usually used in deception research. Several issues were discussed during the interview, such as loyalty amongst Islamic countries, Israel, President George Bush Senior, Western hostages who were used as human shields, and Iraqi planes landing in Iran.

Davis and Hadiks' observations revealed that Hussein used a variety of hand and arm movements, and that he made specific illustrators when discussing specific issues. When discussing Israel, Hussein made a string of short, vertical, intense jabs of the left forearm, some with a fist, some not. This behavioural pattern only emerged when discussing Israel and Zionism. Hussein showed the least restrained behaviour when he spoke about George Bush Senior. Verbally, he gave the impression that he did not care much for Bush when he said: "I talk to people. . .genuine dialogue with people, not with Mr. Bush." Speaking about President Bush Senior, however, was associated with clear movements of the trunk, and a burst of gesture intensity. Davis and Hadiks interpreted this as strong nonverbal evidence of Hussein's intense, personal animosity towards George Bush Senior.

Arnett also discussed the Iraqi planes landing in Iran. Hussein told Arnett that "There isn't one single Islamic country that is not on our side in this battle." He then went on to describe how Iraqi planes may need at times to land in neighbouring countries. He was also asked whether the planes would return to use. There were good reasons to believe that they would not return, but Hussein did not mention this. Throughout this part of the interview, Hussein's behaviour was restricted and controlled. He slowly sat up, and markedly restricted and even stopped his gesticulations. Davis and Hadiks believe that at this particular moment Hussein was fabricating an answer. Unfortunately, the ground truth is not really established in this case. The finding that Hussein showed unique patterns of behaviour when talking about Israel and Bush Senior illustrates that the speaker's involvement with a topic influences his or her behaviour. This should be taken into account when selecting comparable truths.

Ian Huntley On Sunday 4 August 2002, two 10-year-old schoolgirls left their homes in southern England in the evening and vanished without a trace. It was the start of one of the biggest manhunts in British police history, and the search for the girls gripped the nation's attention for the next few weeks. Two hundred journalists went to the village where the girls disappeared and details about the search for the girls were reported on national television news programmes on a daily basis.

Ian Huntley, a 28-year-old school caretaker, was very much present during this search. On the night of the disappearance of the girls he asked police officers whether he could assist them and offered to unlock the school door so that they could search the school. He also told the police that he had spoken to the two girls in front of his house on the evening when they disappeared. They had asked for his girlfriend, Maxine Carr, who had helped out in their class at school. That made Huntley the last known person who had spoken to the girls. Huntley provided an alibi to the police. He said that he was with his girlfriend on Sunday 4 August, something that she confirmed.

Over the next few days, Huntley was hanging around and asked journalists strange questions. For example, he asked one journalist whether they already had found the clothes of the girls. The journalist found this an odd question, because it gave the impression that he knew that the girls were separated from their clothes. She reported this to the police. Meanwhile the police were distracted by alleged sightings of the girls, and by checking the whereabouts of hundreds of known sex offenders in the area.

Journalists asked Huntley to appear on television. Initially he declined, but after six days he changed his mind. However, he only wanted to appear on local television because, according to Huntley, the disappearance of the girls was "not in the national interest". When he was asked in the television programme whether the moment he saw the girls outside his house has been playing on his mind, he replied: "Absolutely, it was very upsetting you know to think that I might have been the last friendly face the girls have spoken to".

Several days after this TV interview Huntley changed his mind again and decided on 14 August to appear on national television. This led to his undoing. People phoned in, including a woman who reported that her daughter was sexually assaulted by Huntley when the daughter was 10 years old. However, the case never came to court. It then became known that Huntley had been investigated before about sex with children. This knowledge, together with evidence that one of the girl's mobile phones was disconnected outside Huntley's home, and the fact that CCTV footage showed that his girlfriend Maxine Carr had been 100 miles away on 4 August and had not been with her boyfriend as she had claimed, made Huntley the main suspect. Shortly after, the police found the girls' clothes in a school storage building Huntley had access to, and his fingerprints were found on a bin bag that covered the clothes. More evidence came to light and Ian Huntley has been found guilty of murdering the two girls.

On the evening of his conviction, BBC television broadcast a programme about this case that included video footage of the police interviews with Huntley and his girlfriend Maxine Carr. To my knowledge, extensive video footage of an interview with a suspect had never been shown on British television before. The programme therefore gave the general public unique insight into the behaviour of a man who lied about murder during a police interview. I noticed three cues to deceit in the video footage: a decrease in hand/finger movements; long pauses;

and word repetitions. In the police interview we come to know Huntley as a man who makes some right-hand illustrators and head scratches with his right hand when he is talking. When he is not making illustrators or head scratches, he makes finger movements with his right hand. However, when he lied (he told the police that the girls walked away towards the library whereas in fact they entered his house) he stopped making these finger movements. The head scratches and illustrators continued. He also included a long pause before telling the interviewer that the girls walked off to the library. A long pause also occurred when he was asked what his girlfriend was doing on Sunday 4 August. He eventually said: "She was doing homework, yes, she was doing homework". I believe that this repetition is the third cue to deceit. It is remarkable, particularly because the second time he mentioned the homework, he makes a positive head nod, but does not really show behaviour meant to convince the interviewer, such as looking the interviewer into the eye or showing positive body direction. It therefore gives the impression that he wanted to convince himself about the veracity of this answer. A repetition also occurred when he was asked whether he had ever had physical contact with the girls. He denies this but also repeats part of the question: "Physical contact? No".

The ground truth is well established in this case. We know that Huntley lied when he said that the girls walked in the direction of the library. We also know that his girlfriend was not at home on that particular Sunday, and that he did have physical contact with the two girls. However, there are problems with the comparable truth. Huntley knew he was a suspect and was probably aware that his answers to the question where the girls went to after he spoke to them and whether he had physical contact with the girls were crucial. These were particularly high-stakes questions.

The murderer A person went missing and was found dead a couple of days later. It was clear that the victim had been murdered. Several independent witnesses told the police that they had seen a man talking to the victim a couple of days before the body was found. On the basis of their descriptions, the police were able to create an artist's impression of the man. After a couple of months, a man was arrested and brought to a police station for interviewing. Apart from the fact that he showed a clear resemblance to the face in the sketch, there were other reasons that led the police to believe that he was involved in the crime.

The police interviewed this man extensively. During the first interview, he was asked what he had done during the day the victim went missing. He described in detail that he went to his work in the morning, visited a friend and a market in the afternoon, and visited a neighbour in the evening.[18] The police checked every single detail the man had provided. Several independent witnesses, including his employer, could confirm his story about his activities during the morning, but no confirmation could be obtained about his alleged activities during the rest of the day. This made the man even more suspicious, and an intensive investigation started. Meanwhile, the man consistently denied having killed the victim, and even claimed that he had never met the victim. After a couple of weeks, substantial evidence was found which made clear that he was the murderer. A hair found in the man's car was confirmed to have come from the victim. In addition, fibres of the cloth in which the dead body was wrapped were found in his car. On the basis of this substantial evidence, the man admitted to having killed the victim. He also gave a detailed description of what had happened. The man was later convicted for murder by a criminal court and sentenced to imprisonment for the remainder of his natural life.

During his confession the murderer did not tell the whole truth either. He told the truth about how he drove from his house to the location where he met the victim, and independent witnesses could confirm this part of the story. It is evident that he lied about how he met the victim. Several independent witnesses claimed to have seen him at a particular location. In addition, an object that belonged to the man (ownership of which he admitted) was found at that location. Despite this substantial evidence, the man continued to deny ever having visited the location. He admitted to having been near that location, but denied ever having actually entered it.

We analysed all the parts of the interview about which we were certain that he was telling the truth or lying (Vrij & Mann, 2001a). Two fragments – one truth and one lie – were derived from the interview before the confession took place. The truth consists of a description of his activities during the morning that the victim went missing. As stated previously, witnesses could confirm this part of the story. This description lasted 61 seconds. The lie, which lasted 67 seconds, consists of a description of his activities during the afternoon and evening of the same day. He told the police about several things he had done in his home town. In reality, he took his car and drove to another city where he met the victim. He killed the victim later on that day.

[18] Although the interview took place several months after the victim went missing, the man was capable of giving a detailed description of his activities on that day. He told the police that he had thought that they would interview him about this case, and had therefore checked his diary to find out what he had been doing on that day. (Even an innocent suspect could have known what day the person went missing, as the media reported extensively on the missing person both during that day and leading up until the body was found.)

Four other fragments – two truths and two lies – were derived from his confession. In the first truthful fragment, which lasted 26 seconds, the man gave a detailed description of how he drove from the exit of the motorway to the place where he met the victim. Witnesses could verify this part of the story. The second truthful fragment, which lasted 27 seconds, was a repetition of the same story. His first lie during his confession was about the time he left a friend's house at his home town the day he killed the victim. In reality, he left the friend's house a couple of hours earlier than he claimed. Witnesses saw him with the victim at the time when he claimed to have been at his friend's place. This is an important lie because he had to account for a couple of hours (the hours when he visited the location which he denied ever visiting). This lie took 16 seconds. The second lie during the confession was about where he met the victim. This lie lasted 32 seconds. As described above, there is compelling evidence that he met the victim at a location that he denied having visited.

Although we have more than 20 hours of videotapes, only a few minutes could be used in this study. For the other parts of the interview the ground truth could not be established. For example, the murderer gave a detailed description of the talks he had with the victim, and how he killed the victim. However, there is no possibility of verifying the truthfulness of this part of his story.

Table 3.2 shows an overview of the man's behaviour. The reported means represent the frequency of occurrence per 100 words (hesitations and errors), length in seconds on a per minute basis (pause duration and gaze aversion), or frequency of occurrence on a per minute basis (other behaviours minus speech rate). Speech rate was measured on a 3-point scale ($1 = \text{slow}$, $3 = \text{fast}$).

A few differences emerged between the truthful and deceptive accounts before the confession. While lying, the murderer made more speech errors, made longer pauses, spoke more slowly, and showed more gaze aversion than when he was telling the truth. This behavioural pattern is typical of someone who has to think hard. Apparently it was more difficult for the man to lie than to tell the truth. It is perhaps surprising that the man apparently had to think hard when he lied. He knew that he was a suspect, and there was enough time for him to prepare a lie. There was also evidence that he had prepared, as he had made false notes in his diary in an effort to dupe the police. Despite this preparation, a possible reason why the man's behaviour still showed some evidence that he had to think hard is that he is not very bright (this is the opinion of the police detectives who interviewed him). There is evidence that suggests that preparation probably does not benefit liars who are not so clever (Ekman & Frank, 1993).

Table 3.2 The behaviour of a murderer during a police interview

	Before confession		During confession	
Behaviour	Truth	Lie	Truth	Lie
Hesitations	11.54	08.04	03.47	03.49
Errors	07.69	14.29	08.33	11.63
Pause duration	10.33	22.84	04.53	15.63
Pause frequency	05.90	06.27	04.53	08.13
Speech rate	02.00	01.00	03.00	02.00
Gaze aversion	26.56	44.33	47.55	12.50
Smiling	00.00	00.00	00.00	00.00
Illustrators	01.97	01.79	06.79	00.00
Self-adaptors	00.00	00.00	00.00	00.00
Hand/finger movements	00.98	01.79	06.79	00.00
Head movements	19.18	17.01	16.98	16.88
Trunk movements	00.00	00.00	00.00	00.00

The reported means represent the frequency of occurrence per 100 words (hesitations and errors), length in seconds on a per minute basis (pause duration and gaze aversion), or frequency of occurrence on a per minute basis (other behaviours minus speech rate). Speech rate was measured on a 3-point scale (1 = slow, 3 = fast).

Several differences between truth telling and lying also emerged in the confession interview. While lying, the murderer made more speech errors, made more and longer pauses in speech, spoke at a slower rate, showed less gaze aversion, and made fewer illustrators and hand/finger movements. Again, the increase in speech errors, the longer and greater number of pauses, and the slower speech can be interpreted as signs that he had to think hard. Also the decrease in movements may be the result of cognitive effort; alternatively, it could be a sign of attempting to appear credible. Maintaining eye contact with the policeman when he was lying may be interpreted as an effort to make a credible impression. The man may have realised that gaze aversion appears suspicious, and, in order to avoid making a dishonest impression, he looked the police detective straight into the eyes. Also, in his confession interview he attempted to convince the police officer that he has never been to the location where the police officer claimed he has been (where he met the victim). People typically look their conversation partners in the eye when they try to persuade them.

This persuasion explanation may also explain the difference in gaze behaviour before the confession (gaze aversion when lying) and during the confession (looking the police officer in the eye when lying). In the interview before his confession, the man was not challenged by the police interviewer, but was just given the opportunity to discuss his whereabouts during a particular day. However, in the interview during his confession the police officer had told him that he did not believe his story about how he met the victim. He may therefore have attempted to persuade the police officer to change his mind about this.

In summary, the behavioural pattern showed by the murderer provides evidence that he experienced cognitive load in particular when he lied. He may also have tried to control his behaviour, but he did not seem to behave nervously. To check this interpretation, we showed 65 police officers, who did not know the man and who knew nothing about the case, the six fragments of the interview that we had analysed and asked them to indicate to what extent they perceived the man to be tense, controlling his behaviour, and having to think hard. We did not tell the police officers when the man was telling the truth and when he was lying. The results indicated that the man particularly gave the impression that he had to think hard when he was lying (Vrij & Mann, 2001a).

Murderers, arsonists, and rapists As part of her PhD project, Samantha Mann analysed the behavioural responses of 13 male and three female suspects while they told the truth and lied during their police interviews (Mann *et al.*, 2002; Vrij & Mann, 2003a, b, 2005). The police interviews were videotaped, and the tapes were made available for detailed coding of the suspects' behavioural responses. The suspects were interviewed in connection with serious crimes such as murder, rape, and arson and were facing long custodial sentences if found guilty. In this study the ground truth was established in all 16 cases and fair comparisons between the suspects' truthful and deceptive behaviour were made. Regarding the ground truth, clips of video footage were selected where other sources (reliable witness statements and forensic evidence) provided evidence that the suspect told the truth or lied. In addition, for each suspect, truths and lies were chosen which were as comparable as possible in nature. For example, a suspect who gave a detailed description about how he had assisted in killing a person (truth) later denied any involvement in the crime (lie). Forensic evidence indisputably supported his original version.

Results revealed that, compared to when they told the truth, the suspects exhibited longer pauses and fewer eye blinks (Mann *et al.*, 2002). When the data were analysed for the 13 male suspects only a third cue to deception occurred: liars made fewer hand/finger movements than truth tellers (Vrij & Mann, 2003a, b, 2005). Indicators of being tense (such as fidgeting and gaze aversion) did not emerge. These results suggest that the suspects' cues to deception were more likely the result of increased cognitive demand or attempted behavioural control than of nervousness. The strongest evidence for this was the reduction in eye blinks during deception. Research has shown that negative emotional arousal results in an increase in eye blinking (Chiba, 1985; Harrigan & O'Connell, 1996; Tecce, 1992), whereas increased cognitive load results in a decrease in eye blinking (Bagley & Manelis, 1979; Bauer, Goldstein, & Stern, 1987; Bauer, Strock, Goldstein, Stern, & Walrath, 1985; Holland & Tarlow, 1972, 1975; Wallbott & Scherer, 1991).

To check our assumption that the suspects' cues to deception were more likely the result of increased cognitive demand or attempted behavioural control than nervousness we conducted a follow-up study (Mann & Vrij, 2006). We showed police officers a selection of the truthful and deceptive clips of Mann *et al.*'s (2002) study. After each fragment the officers were asked to indicate to what extent the suspect (i) appeared tense, (ii) gave the impression that he or she had to think hard, and (iii) gave the impression that he or she was controlling him/herself. Results revealed that the suspects appeared to be thinking harder when they lied than when they told the truth. They also appeared to be trying to control themselves more when they lied than when they told the truth. However, the suspects appeared tenser when they told the truth than when they lied.

Martha Davis and her colleagues analysed videotaped statements of 28 criminal suspects (Davis *et al.*, 2005). The suspects gave their statements to assistant district attorneys after they had been interrogated by the police. Ground truth was well established in this study via laboratory evidence, crime scene analysis, or witness and suspect accounts. Interestingly, the researchers also scored the "incriminating potential" (IP) of each utterance. IP was negligible if the utterance was about issues such as nicknames, but very high if the utterance was directly related to the crime. Four IP categories were defined. IP is related to comparable truth, and, ideally, low IP lies should be compared with low IP truths, and high IP lies should be compared to high IP truths. Unfortunately, there were no low IP lies. Davis *et al.* found that suspects did not lie randomly, but that the proportion of lies increased when IP increased. Three cues to deception emerged: word and phrase repetitions, stammers, and protracted headshakes. All three cues occurred more often in deceptive utterances. Although they examined movements of the arms, these were not related to deception. However, they scored the overall category "gestures" and, for example, did not study specific hand/finger movements.

Summary Despite problems with ground truth or comparable truth, a trend emerged from these real-life observations. People's expectations that liars will show nervous behaviours, particularly gaze aversion and fidgeting, were not supported. Rather, liars gave the impression that they had to think hard or that they tried to control themselves. This results in cues such as word and phrase repetitions, long pauses, or a decrease in hand/finger movements.

There are several explanations why those people did not display clear patterns of nervous behaviours. For example, this section contained analyses of criminals and politicians, and such individuals may be different from average people. Perhaps the average person does display clear signs of nervousness when telling high-stakes lies. Although I cannot rule out this explanation, I don't think it holds true. I think that many more people do not exhibit signs of nervousness when telling high-stakes lies. I believe that this is because they also experience high cognitive load when telling such wicked lies. As mentioned above, cognitive load is associated with activity in higher areas of the brain which, in turn, inhibits fidgeting (e.g., signs of nervousness). In addition, there is evidence that when liars have to think hard, their cognitive effort momentarily suppresses their tonic (physiological) arousal (Leal, Vrij, Fisher, & van Hooff, 2006). When someone's physiological arousal decreases, signs of nervousness are less likely to occur.

Interview style

Finally, it could be that the cues to deception that arise depend on the way a person is interviewed. There is virtually no research regarding whether interview style has an effect on nonverbal cues to deception. Exceptions are perhaps the interactive studies that have been carried out as part of the Interpersonal Deception Theory approach. These have been discussed above. However, in these studies different interview styles have never been systematically manipulated and, as a result, these studies do not answer whether different interview styles result in different nonverbal cues to deception. I believe this is a shortcoming, because I think that interview style does affect cues to deception, as I demonstrated in my own study (Vrij, 2006a).

In their analysis of audiotaped interviews with suspects in England and Wales, Moston and Engelberg (1993) observed that the police commonly use two types of interview styles: information gathering and accusatory. In the information-gathering style, interviewers request suspects to give detailed statements about their activities through open questions (e.g., "What did you do between 3pm and 4pm?"; "You just mentioned that you went to the gym last night; who else was there?"). By comparison, in the accusatory style, interviewers confront suspects with accusations (e.g., "Your reactions make me think that you are hiding something from me"). (See also Vrij, 2003, and Hartwig, Granhag, & Vrij, 2005.)

Following Moston and Engelberg's (1993) analysis, truth tellers and liars in my experiment were interviewed in three different phases

(Vrij, 2006a). The interview started with an information-gathering style (Phase 1, "Please tell me what happened when you were in that room just now"), which then developed into an accusatory style (Phase 2, "Your reactions make me think that you are hiding something from me"), and finally transformed back into an information-gathering style (Phase 3, "Please tell me again what happened when you were in that room just now"). When truth tellers and liars were compared, only one cue emerged in the accusatory phase: liars smiled less than truth tellers. In contrast, two cues to deception emerged in each of the two information-gathering phases: in both phases liars smiled less and made fewer hand and finger movements than truth tellers. Perhaps the difference between information-gathering and accusatory interview styles was caused by the length of the statements. Answers to information-gathering questions (e.g., "What did you do this afternoon between 3pm and 4pm?") typically result in longer statements than answers to accusatory questions (e.g., "I think that you did take the money"). Longer statements give more opportunities to display nonverbal cues to deception, and, indeed, longer statements reveal more nonverbal cues to deception than shorter statements (DePaulo, Lindsay *et al.*, 2003).

Additionally, a *context overshadowing effect* could explain why the accusatory phase of the interview resulted in fewer nonverbal cues to deceit than the two information-gathering phases (Vrij, 2006a). The context overshadowing effect is based on the assumption that being accused of wrongdoing (i.e., accusatory interview style) is likely to affect the behaviour of both liars *and* truth tellers. Both liars and truth tellers may experience negative emotions when they realise the consequences of not being believed (Bond & Fahey, 1987; Ekman, 1985/2001); and may try to convince the interviewer that they are innocent by attempting to exhibit behaviours that they believe come across as persuasive or truthful (Buller & Burgoon, 1996; Edinger & Patterson, 1983; Hocking & Leathers, 1980). As a result, the accusation has a stronger impact on someone's nonverbal behaviour than the act of lying, and, consequently, differences in nonverbal behaviour between truth tellers and liars diminish when being accused.

CONCLUSION

How do people behave when they lie? Unfortunately, it is not possible to provide a simple answer to this question. There is no nonverbal behaviour that is uniquely associated with deception, neither is there a theoretical explanation why such cues akin to Pinocchio's growing nose would exist. The behaviours liars may show depend on many factors as this chapter demonstrates. Despite the absence of a simple answer, we now have a hint about which strategy to use to identify nonverbal cues to deception.

The chapter demonstrated that examining a cluster of nonverbal cues gives more valuable information about deception than examining each cue individually. In all studies where researchers investigated a cluster of cues, at least 70% of truths and lies could be correctly classified on the basis of such clusters. Thus, the first guideline is to always examine a cluster of nonverbal cues rather than individual cues. This raises the question of which cluster of cues someone should investigate. There is no straightforward answer to this, because it probably depends on the individual liar (e.g., interpersonal differences) and on the circumstances under which the lie is told (e.g., intrapersonal differences). Therefore, and this is the second guideline, the observer needs to control for such individual differences and circumstances when examining cues to deceit, by examining *changes* in behaviour shown by an individual under *similar circumstances*. I gave a real-life example of a murderer who, when he described his activities during a particular day, showed different behaviours when he described his activities during the morning compared to when he described his activities during the afternoon and evening (Vrij & Mann, 2001a).

An observer should be careful with interpreting such changes. They indicate that something is going on, but not necessarily that the person is lying, because changes in behaviour can be caused by factors other than deception. In that respect, identifying nonverbal cues to deceit is a two-step approach. The first step is noticing changes in behaviour, and the second step is interpreting what these changes actually mean. For the second step, the observer often needs to investigate further, for example, by intensifying the search for evidence. This strategy worked in the case of the murderer. Intensive investigations resulted in finding a hair of the victim in the murderer's car and a piece of the cloth in which the victim's body was wrapped up, which subsequently broke his determination in continuing to deny the crime.

Alternatively, the observer can decide to further interview the alleged liar. In that respect, this chapter revealed that an information-gathering interview style is more helpful than an accusatory style because it reveals more nonverbal cues to deceit. Using an information-gathering interview style as starting point, I believe that interviewers could take a proactive role by developing interview protocols that magnify the differences between truth tellers and liars. In Chapter 15 I will discuss in detail what I have in mind.

Appendix 3.1 Nonverbal cues to deception

	Vocal Cues						
	Hesitations	Speech errors	High–pitch voice	Speech rate	Latency period	Pause durations	Pause frequency
Abe *et al.* (2006)					–		
Allen *et al.* (1992)					>		
Anolli & Ciceri (1997)		>	>	–		–	>
Anolli *et al.* (2003)	–	>				–	>
Bello (2006)	–	>					>
Bond *et al.* (1985)	–	–					
Bond *et al.* (1990)	>	–	–				–
Buller & Aune (1987)			–	–			
Buller *et al.* (1989)		<			–		<
Buller *et al.* (1994a)	>	–			<		
Cody & O'Hair (1983)					<		
Cody *et al.* (1984, immediate response)		–		–	>	>	–
Cody *et al.* (1984, delayed response)		>		–	–	>	–
Cody *et al.* (1989)	–	>		–	–		–
Davatzikos *et al.* (2005)					<		
Davis *et al.* (2005)	<	>		–			–
DePaulo *et al.* (1982b)	–	>		–			
DePaulo, P. J. & DePaulo (1989)	–	–					
DeTurck *et al.* (1985)		>			>	>	
Dulaney (1982)		–			<		

Ebesu & Miller (1994)	–			<		–
Ekman (1988)			>		<	
Ekman *et al.* (1976)			>			
Ekman *et al.* (1985)					>	
Ekman *et al.* (1991)			>			
Engelhard *et al.* (2003)					–	
Farrow *et al.* (2003)					>	
Feeley & deTurck (1998)	>	>		–	–	–
Fiedler & Walka (1993)	>		–	<		
Fukuda (2001)					>	
Ganis *et al.* (2003)					–	
Gozna & Babooram (2004, seated interviews)	–		–	–	<	
Gozna & Babooram (2004, standing interviews)	<		–	>	<	
Granhag & Strömwall (2002)	–			–	–	<
Greene *et al.* (1985)					>	
Gregg (2007, exp. 1)					>	
Gregg (2007, exp. 2)					>	
Harrison *et al.* (1978)					>	
Heilveil & Muehleman (1981)		>			>	–
Hocking & Leathers (1980)	–	>		–	–	–
Höfer *et al.* (1993)	–	<		–	<	
Johnson, R. *et al.* (2003)					>	
Kalma *et al.* (1996)					<	<
Klaver *et al.* (2007)	–	–		–		–
Knapp *et al.* (1974)	–	–				
Köhnken (1985)	–	>		<		

(Continued)

Appendix 3.1 *(Continued)*

	Vocal Cues						
	Hesitations	Speech errors	High–pitch voice	Speech rate	Latency period	Pause durations	Pause frequency
Kraut (1978, exp. 1)	–	–			>		
Kraut (1978, exp. 2)					<		
Kraut & Poe (1980)		–			–		
Langleben *et al.* (2005)					>		
Levine *et al.* (2005)		<		–	<		<
Mak & Lee (2006)					>		
Mann *et al.* (2002)	–	–				>	
Matarazzo *et al.* (1970)					–		
Mehrabian (1971, exp.1)		–					
Mehrabian (1971, exp. 2)				–			
Mehrabian (1971, exp. 3)				>			
Meharabian (1972)		>		–			
Miller *et al.* (1983)	–	–			–		–
Motley (1974)			–				
Nunez *et al.* (2005)					>		
O'Hair *et al.* (1981)					<		
Parliament & Yarmey (2002)					<		
Porter & Yuille (1996)	–						
Porter, Doucette *et al.* (in press)	–			–			–
Riggio & Friedman (1983)	–						
Rockwell *et al.* (1997)			–	<	>		
Rosenfeld, Biroschak *et al.* (2006)					–		
Seymour *et al.* (2000)					>		
Spence *et al.* (2001)					>		

Spence *et al.* (2003)					>		
Stiff & Miller (1986)	–				–	–	
Stiff *et al.* (1994)					–		
Streeter *et al.* (1977)			>				
Strömwall *et al.* (2006)	–			–		–	–
Vendemia *et al.* (2005a, exp. 1)					>		
Vendemia *et al.* (2005a, exp. 2)					>		
Vendemia *et al.* (2005a, exp. 3)					>		
Vendemia *et al.* (2005b)					>		
Verschuere *et al.* (2004a)					–		
Verschuere *et al.* (2004b, exp. 1)					>		
Verschuere *et al.* (2004b, exp. 2)					>		
Verschuere *et al.* (2004b, exp. 3)					>		
Vrij (1995)	<	–	>	–			
Vrij (2006a, information–gathering interview)	–	–					
Vrij (2006a, accusatory interview)	–	–					
Vrij & Winkel (1991)	<	–	–	>			
Vrij & Heaven (1999, difficult response)	>	>					
Vrij & Heaven (1999, easy response)	<	–					

(Continued)

Appendix 3.1 *(Continued)*

	Vocal Cues						
	Hesitations	Speech errors	High–pitch voice	Speech rate	Latency period	Pause durations	Pause frequency
Vrij, Edward *et al.* (2000)	>	–		–	>		
Vrij & Mann (2001a)	–	>		<		>	–
Vrij *et al.* (2001a)	–	–		–	>		
Vrij *et al.* (2001c)	>	–		>	>		
Vrij *et al.* (2004a, adults)	–	–		–	–	–	
Vrij *et al.* (2006a)					–		
Vrij, Mann *et al.* (in press, normal order)	–	–		–	–	–	
Vrij, Mann *et al.* (in press, reverse order)	–	>		<	–	–	
Walczyk *et al.* (2003)					>		
Walczyk *et al.* (2005, exp. 1)					>		
Walczyk *et al.* (2005, exp. 2					>		
Winkel & Vrij (1985)	–	>					
Zhou & Zang (2006)					>		<
Summary	–	–	>	–	>	>	–
DePaulo, Lindsay *et al.* (2003)	.04	.00	**.21**	.07	.02		

< cue is shown less by liars than by truth tellers
> cue is shown more by liars than by truth tellers
– no differences was found between liars and truth tellers
Empty cells mean that the cue was not investigated

Appendix 3.1 *(Continued)*

	Visual Cues									
	Gaze	Smile	Self adaptors	Illustrators	Hand/finger	Leg/foot	Trunk	Head	Shifting position	Eye blink
Akehurst & Vrij (1999)			–	–	<	–	–			
Bond *et al.* (1985)	>	–	–	>				–	–	
Bond *et al.* (1990)	–	–	–	–				–		–
Buller & Aune (1987)	<	–	<	–		–	–	–	–	
Buller *et al.* (1989)	<	–	–	–				–		
Buller *et al.* (1994a)		–						–		
Buller *et al.* (1994b)								<		
Burns & Kintz (1976)	–									
Caso *et al.* (2006a)			<	–						
Caso *et al.* (2006b)				<	<	–				
Chiba (1985)										>
Cody & O'Hair (1983)	–	–	–	<		<			>	
Cody *et al.* (1989)	–	–	–	–	–			–		–
Craig *et al.* (1991)										<
Davis & Hadiks (1985)				<						
Davis *et al.* (2005)				–				>		
Dente *et al.* (2006)					–					
DePaulo, P. J. & DePaulo (1989)	–		–					–		
deTurck & Miller (1985)	–		>	>		–				–
Donaghy & Dooley (1994)								–		
Ebesu & Miller (1994)			–	–					>	–

(Continued)

Appendix 3.1 *(Continued)*

	Visual Cues									
	Gaze	Smile	Self adaptors	Illustrators	Hand/finger	Leg/foot	Trunk	Head	Shifting position	Eye blink
Ekman (1988)		–	<	<		<				
Ekman & Friesen (1972)			>	<						
Ekman *et al.* (1976)				<						
Ekman *et al.* (1988)		–								
Ekman *et al.* (1990)		–								
Ekman *et al.* (1991)		–		<						
Feeley & deTurck (1998)	–	–	–	–		–			–	
Fiedler & Walka (1993)		>	–					<		
Fukuda (2001)										<
Galin *et al.* (1993)										>
Gozna & Babooram (2004, seated interviews)	–	–	–	–	<	<	–	>	–	
Gozna & Babooram (2004, standing interviews)	–	>	–	–	–	–	–	>	–	
Granhag & Strömwall (2002)	>	<	<	–			–	–		
Greene *et al.* (1985)	–	–	–	<	<	<		<	–	
Griffin & Oppenheimer (2006)	>									
Gross & Levenson (1993)	–	–	<				<			>
Hadjistavropoulos & Craig (1994)										–
Hadjistavropoulos *et al.* (1996)										–

Heilveil & Muehleman (1981)	–	–	–					–		
Hess & Kleck (1994)	–									>
Hill & Craig (2002)										–
Hocking & Leathers (1980)	<	–	–	–		<		–	–	–
Höfer *et al.* (1993)	–	–	–	<		<	–	–		
Kalma *et al.* (1996)	<	–		<		>				
Kennedy & Coe (1994)	–	–								–
Klaver *et al.* (2007)		–	–	–	–	–		>	–	–
Knapp *et al.* (1974)	<	–	>	–		–		–		
Kraut (1978, exp. 1)		–	–				–		–	
Kraut & Poe (1980)	–	–	–	–			–		–	
Leal *et al.* (2006)										<
Levine *et al.* (2005)	–			–		<				
Mann *et al.* (2002)	–		–	–	–					<
Matarazzo *et al.* (1970)	–									
McClintock & Hunt (1975)	–	–	>	–			>		>	
Mehrabian (1971, exp. 1)						<				
Mehrabian (1971, exp. 2)				–				–		
Mehrabian (1971, exp. 3)			–				–	–		

(Continued)

Appendix 3.1 *(Continued)*

	Visual Cues									
	Gaze	Smile	Self adaptors	Illustrators	Hand/finger	Leg/foot	Trunk	Head	Shifting position	Eye blink
Mehrabian (1972)				–		<		–		
Miller *et al.* (1983)	<		–	–	–	–				
O'Hair *et al.* (1981)	–	<	>	–		–		>	–	
Porter, Doucette *et al.* (in press)		–	>	>				–		
Riggio & Friedman (1983)	>	–	–	–		–		–	–	
Schneider & Kintz (1977)						–				
Sitton & Griffin (1981)	>									
Stiff & Miller (1986)		–	–	–		–			–	–
Strömwall *et al.* (2006)	–	–	–	–			–	–		
Vrij (1995)	–	–	–	–	<	<	–	–	–	
Vrij (2006, information-gathering interview)	–	<	–	–	<	–				
Vrij (2006, accusatory interview	–	<	–	–	–					
Vrij & Winkel (1991)	–	–	–	–	<	<	–	–	–	
Vrij, Semin *et al.* (1996)			–	–	<	<		–	–	
Vrij, Akehurst *et al.* (1997)				–	<					

Vrij, Edward *et al.* (2000)	–	–	–	<	<	–				
Vrij & Mann (2001a)	–	–	–	<	<		–	–		
Vrij *et al.* (2001a)	–	<	–	<	<	<				
Vrij *et al.* (2001c)	–	–	–	<	–	–				
Vrij *et al.* (2004a, adults)	–		–	–	<	–				
Vrij *et al.* (2006a)	–		–	–		<			<	
Vrij, Mann *et al.* (in press, normal order)	–		–	–	<	–				
Vrij, Mann *et al.* (in press, reverse order)	–		–	–	–	>				>
Zuckerman *et al.* (1979)		<								
Summary	–	–	–	<	<	<	–	–	–	–
DePaulo, Lindsay *et al.* (2003)	.03	.00	–.01	**–.14**	**–.36**	–.09		–.02		.07

< cue is shown less by liars than by truth tellers
> cue is shown more by liars than by truth tellers
– no differences was found between liars and truth tellers
Empty cells mean that the cue was not investigated
DePaulo, Lindsay *et al.* (2003) did not measure trunk movements and shifting positions separately

CHAPTER 4

Individual Verbal Cues to Deception

This chapter is the first of five chapters that examine verbal cues to deception. The verbal cues discussed in Chapters 7, 8, 9, and 10 form part of existing verbal lie detection tools used by professional lie catchers and scholars. I summarise in the present chapter verbal cues that are not part of such existing tools, and call these cues "individual cues". Box 4.1 gives descriptions of these individual cues.

Box 4.1 A description of the verbal behaviours

1. *Negative statements*: statements indicating aversion towards an object, person or opinion, such as denials, disparaging statements and statements indicating a negative mood.
2. *Generalising terms*: the use of words such as "always", "never", "nobody" or "everybody".
3. *Self-references*: the use of words referring to the speaker himself/herself such as "I", "me" or "mine".
4. *Immedicacy*: speaker's response sounds direct, relevant, and clear rather than indirect, distancing, or evasive.
5. *Response length*: length of the answer or number of spoken words.

6. *Plausible answers*: statements that make sense and that sound credible and reasonable.
7. *Lexical diversity*: number of different words in a statement divided by the total number of words used in that statement.
8. *Consistencies*: number of the same details repeated over two statements by the same speaker or mentioned by two different speakers.
9. *Contradictions*: contradictions within a statement or between two statements.

Speech content can reveal deception, particularly when the liar says something that the observer knows is not true. A straightforward comparison between the facts and statements from the alleged liar can reveal deceit beyond doubt in such a situation. In a real-life case, Jeffrey Archer, an author and former English Conservative politician, met with three journalists in his hotel room during a political party conference. When he answered his phone, he asked the three journalists to leave because, Archer claimed, it was the Prime Minister who rang. Another politician who saw the three journalists pacing up and down the corridor asked them what they were doing. He immediately realised that Archer had lied to the journalists and could not be speaking to the Prime Minister on the phone, because he knew that the Prime Minister was sitting on the conference platform at that very moment (*The Independent*, 20 July 2001, p. 3). Archer had told a lie that was discovered because his story contradicted the facts.

Another liar caught out by factual evidence was Jonathan Aitken, a former English Conservative politician who served under Prime Minister John Major, first as Defence Procurement Minister and then as First Secretary to the Treasury. While serving as Minister for Defence Procurement, Aitken spent some time in the Ritz hotel in Paris. The British newspaper *The Guardian* and the television programme *World in Action* reported that the hotel bill was paid by a businessman. Aitken strongly denied this. He claimed instead that his wife had paid the bill, and he sued the newspaper and television station for libel. The libel trial collapsed because the accused parties revealed evidence showing that the bill was paid by a Saudi potential arms customer rather than by his wife. Aitken faced a jury trial and admitted charges of perjury and perverting the course of justice. He was sent to prison.[1]

[1] This is a remarkable story. I suspect that Aitken was aware that the lie about the whereabouts of his wife could be discovered. By suing the newspaper and television programme he probably increased the likelihood of his getting caught, by encouraging the accused parties to search and reveal evidence that he was lying.

In many cases the observer will not know any facts, and in such cases lies cannot be discovered by comparing a statement with factual evidence. It could also be that the observer knows some facts, but that s/he cannot conclude that someone is lying by comparing those facts with the statement. A suspect may admit that he entered a woman's apartment, but may deny that this was against her will. In this case, physical evidence that the man was in the woman's flat does not shed light on the veracity of his denial.

The question arises whether aspects of the speech content reveal deception. For example, do liars tend to say certain things, or do they tend to avoid saying certain things? This question will be discussed in the five chapters on verbal cues to deception. As I will demonstrate, a verbal cue uniquely related to deception, akin to Pinocchio's growing nose, does not exist. However, some verbal cues can be viewed as weak diagnostic indicators of deceit, including some of the individual cues discussed in this chapter.

THEORETICAL PERSPECTIVES

Some of the aspects that affect nonverbal behaviour during deception (e.g., emotions, cognitive effort, attempted control, lack of embracement, Chapter 3) may also influence speech content during deception.

Emotions

I already explained in Chapter 3 that liars sometimes feel guilty or that they are sometimes afraid of getting caught out. Guilt and fear are both negative emotions and may leak through verbally by liars making negative comments. President Nixon for instance said "I am not a crook" instead of "I am an honest man" in the Watergate scandal (DePaulo, 1998, cited in *The Sunday Times*, 24 May 1998, p. 14). Other examples of negative comments is the use of words that reflect negative affect, such as "uncomfortable", "hate", "useless", "dislike" and so on. A boy who secretly took some sweets may reveal his deceit by falsely claiming that he dislikes those sweets.

Another consequence of experiencing negative emotions could be that liars do not want to associate themselves with their lies and therefore may be inclined to give answers that contain general terms or that do not explicitly refer to themselves. For example, a possible answer to the question "Do you smoke?" could be: "Nobody smokes in this house". The tendency to disassociate themselves from a statement could also result in statements that are indirect, distancing, and evasive (e.g., non-immediate). When President Clinton was asked in the Paula Jones case whether Monica Lewinsky was lying when she told someone that she had a sexual affair with him that started in November 1995, Clinton answered: "It's certainly not the truth. It will not be the truth" (*The Independent*, 30 July 1998, p. 14). Disassociation with the statement could finally lead to providing short answers.

Cognitive Effort

Fabricating a lie can be difficult, particularly if the liar has no opportunity to prepare him/herself. When liars experience difficulty in fabricating a lie, their speech content may give the lie away. For example, a lack of creativity in fabricating a lie could lead to a shortage of specific detail, and this may result in using more generalising terms, fewer self-references, or shorter statements. Also, liars may not succeed in fabricating a convincing lie and the lie may therefore sound implausible. Furthermore, liars may not be able to think of something that will assure them of getting away with the lie and therefore may decide to avoid getting into details, and say something evasive, unclear, or irrelevant (e.g., non-immediate). Another cognitively demanding aspect of lying is that liars need to remember what they have said before so that they are able to repeat the information when they are asked to tell their stories again. Memory failures may lead to liars contradicting themselves or being inconsistent (i.e., not the same details are repeated over time).

Attempted Control

Liars may also actively attempt to avoid producing statements that are self-incriminating. They then face a dilemma. The safest strategy is to keep silent, because when liars do not say anything, their speech cannot give their lies away. However, liars may well refrain from remaining quiet because they may realise that this will appear suspicious. The best solution therefore could be to give a statement that provides the observer with as little lie-catching opportunities as possible. Liars could achieve this by using general, non-specific language. They could further achieve this by not referring to themselves, by providing short statements that do not provide too much detail or by giving irrelevant "non-immediate" information as a substitute for information that the deceiver does not wish to supply. Regarding the latter, the boy who wants to conceal that he went to the cinema with his friend, could focus on telling his parents, who assume that he went to the zoo, what he discussed with his friend when they ask him to describe his visit to the zoo.

Attempting to avoid making self-incriminating remarks may also lead to speaking with greater caution. This may result in a liar using a greater number of uncommon words that will lead to greater lexical diversity (Carpenter, 1981). Theoretically, however, an opposite effect, lying leads to lesser lexical diversity, may also occur. When people try hard to get their message across they sometimes repeat the information they provide (Aune, Levine, Park, Asada, & Banas, 2005). Repetition leads to lower lexical diversity. Moreover, increased motivation to appear credible typically leads to more stereotypical language, and this is also characterised by lower lexical diversity (Hollien, 1990; Osgood, 1960).

Finally, attempting to control speech content may result in highly consistent story telling. That is, a liar could prepare the story he or she is going to tell in advance and will stick by this rehearsed and scripted story every time he or she is asked to repeat it. In contrast, truth tellers will most often recall stories from their memories. This may result in changes each time the story is retold, because they may forget to tell information the second time that they told the first time or remember to include something the second time that they forgot the first time.

Lack of Embracement

Many of the verbal cues could also occur because liars lack conviction when making their statements. Liars may lack conviction because they feel uncomfortable when they lie due to feeling guilt or experiencing fear. I already discussed above how guilt and fear could influence verbal cues to deception.

Liars may also lack conviction because they have not personally experienced the claims they make. This lack of personal experience may result in using more generalising terms, fewer self-references, or providing shorter answers.

THE SPEECH CONTENT OF A LIAR

Individual verbal cues to deception are investigated in the same way that nonverbal cues to deception are examined. Typically, truth tellers and liars participate in a laboratory study and produce statements for the sake of the experiment. Those statements are then transcribed and the individual verbal cues are coded. I have already discussed such studies in more detail in Chapter 3.

Appendix 4.1 summarises the findings of each of the 69 studies which together make up all of the studies published in English that I am aware of, where verbal behaviour of adult truth tellers and liars has been compared. In all these studies the truth tellers and liars were adults. For each verbal cue I have indicated its relationship with deception. A "<" sign indicates that the verbal cue appeared less often in liars' statements than in truth tellers' statements; a ">" sign means that the verbal cue was expressed more often by liars than by truth tellers; and a "–" sign means that no difference was found between liars and truth tellers. Empty cells mean that the verbal cue was not investigated in that particular study. At the bottom of Appendix 4.1 I have included a summary of the results of DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper's (2003) quantitative meta-analysis of cues to deception. The numbers in the bottom cells of Appendix 4.1 represent the effect sizes, *d*, whereby effect sizes of .20, .50, and .80 should be interpreted as small, medium, and large effects, respectively (Cohen, 1977). Negative *d*s indicate that the verbal cue was expressed less often by liars than by truth tellers, whereas positive *d*s indicate that the verbal cue was expressed more often by liars than by truth tellers. Only the *d*s printed in bold were significant, that is, indicate a real difference between truth tellers and liars.

First of all, Appendix 4.1 shows that no individual verbal cue obtained the same outcome in every single study in which it was examined. This means that an individual verbal diagnostic cue, similar to Pinocchio's growing nose, does not exist. However, a trend emerged for negative statements, generalising items, self-references, and plausibility. For all these cues where a difference was found between truth tellers and liars, the difference that emerged was predicted by theory as described above. That is, the statements that liars produce include more negative comments, more general items, and less self-references, and sound less plausible than the statements that truth tellers produce. DePaulo, Lindsay *et al.*'s (2003) meta-analysis revealed that the effects for negative comments and plausibility are real, albeit somewhat weak. Including negative comments could be the result of liars experiencing negative emotions that leak through to their speech, and providing less plausible answers could be the result of liars finding it too difficult to come up with a plausible answer. The trends regarding generalised statements and self-references did not appear to exist in DePaulo, Lindsay *et al.*'s meta-analysis, probably because in the majority of studies no relationship was found between these two cues and deception.

Appendix 4.1 also shows that several studies found that liars are less immediate than truth tellers. Liars may give non-immediate answers because they feel negative about their lie and want to dissociate themselves from it, because they cannot quickly think of an appropriate answer, or because they do not wish to reveal relevant information. This idea received rather strong support in DePaulo, Lindsay *et al.*'s (2003) meta-analysis, $d = -.55$. However, most of the findings regarding the immediacy effect were obtained by one group of researchers. I believe that a cue receives stronger support if different researchers are able to reproduce the same finding. Other researchers were unable to replicate the finding that liars are less immediate in their responses. In fact, we found the opposite effect, and because our study was conducted and published after DePaulo, Lindsay *et al.* published their meta-analysis, our finding was not incorporated in that meta-analysis. In all, I think that the support for immediacy as a diagnostic cue to deception is less strong than the d-value suggests.

The findings related to response length are erratic. Most researchers, however, have found that liars provide shorter answers than truth tellers, as is predicted by theory. In their meta-analysis Sporer and Schwandt (2006a) found just such a relationship between lying and response length, albeit small, $d = -.11$. Liars may produce shorter answers because they distance themselves from their statements, because they find lying cognitively difficult, or because they deliberately give short answers so that they do not provide observers with many lie-catching opportunities.

No clear pattern emerged regarding lexical diversity, and it is therefore safest to conclude that there is no clear link between lexical diversity and deception. This is not surprising because, as discussed above, the attempted control approach could predict either an increase or a decrease of lexical diversity as a consequence of lying.

Consistency, the number of details that are repeated in two different statements about the same subject, does not appear to be related to deception when successive statements from the same person are compared. However, a trend emerges when statements from two different people about the same event are taken into account: lying pairs show more consistency than truth telling pairs. This is easy to explain. Two liars may discuss amongst themselves what they are going to say and subsequently recall this agreed scripted story. Their stories are then likely to be similar. In contrast, two truth tellers both use their memory of the event as the source of their recall. Their memories may differ slightly, or the truth tellers may differ in which details of the event are worth recalling. Less consistent recall is then the result (e.g., "repeat versus reconstruct hypothesis", Granhag, Strömwall, & Jonsson, 2003).[2] Finally, Appendix 4.1 shows no evidence that contradictions are related to deception.

[2] Judging consistency within a statement may be more difficult than it initially appears. Observers often disagree amongst each other about whether a statement is consistent with a previous statement or not (Granhag & Strömwall, 2001a, b).

DePaulo, Lindsay *et al.* (2003) looked at almost 30 more individual verbal cues to deception studies than I have included in Appendix 4.1, but most of them were measured incidentally. Only seven cues were measured more frequently but none of these cues were related to deception. Those cues include the use of group references (e.g., using words such as "we", "us" or "ours"), other references (e.g., using words such as "he", "she", "they" or "them"), tentative constructions (e.g., using words such as "may", "might", "could", "I think" or "I guess") and ritualised speech (e.g., using words such as "you know", "well", "really" or "I mean").

Correct Classifications of Truth Tellers and Liars on the Basis of Individual Verbal Cues

Some researchers have examined the extent to which truth tellers and liars can be correctly classified on the basis of individual verbal cues. In all these studies, researchers examined a cluster of verbal cues rather than each cue individually. Looking at individual cues is searching for the verbal equivalence of Pinocchio's growing nose, which does not exist. Conversely, examining a cluster of verbal cues has yielded successful classifications in 67% to 80% of truth tellers and liars (Bond & Lee, 2005; Colwell, Hiscock, & Memon, 2002; Newman, Pennebaker, Berry, & Richards, 2003; Zhou, Burgoon, Twitchell, Qin, & Nunamaker, 2004b).[3] However, different researchers examined different verbal clusters. It is unknown to what extent a certain cluster that works in one situation or one group of participants also works in another situation or with another group of participants.

SITUATIONAL AND INDIVIDUAL DIFFERENCES

Are any of the individual verbal cues discussed in this chapter stronger cues to deception under certain circumstances or in certain individuals? Unfortunately, I cannot give a definite answer to this question because it has not been examined often enough. In their meta-analysis, DePaulo, Lindsay *et al.* (2003) examined whether cues to deception differ when people are highly motivated or not so motivated to appear credible, or when they have or do not have the opportunity to prepare

[3] In some of these studies the frequency of occurrence of the verbal cues was automatically generated via computer programs, saving the time-consuming exercise of human coding. Automatic coding, however, is necessarily restricted. For example, it is not possible to judge the plausibility of a statement via computerised coding.

their answers. Only the response length was measured often enough in research to be included in those analyses. DePaulo, Lindsay *et al.* found that motivation and planning had no effect on response length. However, Sporer and Schwandt (2006b) found in their meta-analysis that highly motivated liars gave shorter answers than highly motivated truth tellers, whereas no differences in response length emerged between truth tellers and liars who were not so motivated. Also, liars who had some time to prepare themselves tended to talk less than truth tellers who had been given some preparation time, whereas no differences between truth tellers and liars emerged when they had a short preparation time.

Regarding individual differences, I am not aware of studies examining gender or ethnic (cultural) differences in individual verbal cues to deception.[4] Observational data in daily life settings has revealed that four year olds' lies typically take the form of one-word responses rather than the more sophisticated elaborations of older children and adults (Bussey, 1992). Obviously, many of the cues discussed in this chapter do not appear in one-word answers, and, consequently, those individual cues to deception cannot be examined in very young children. Over the years, children will become better verbal liars (Talwar, Murphy, & Lee, 2007), particularly when they start to consider the listener's mental state, as they will do by the age of four (Bussey, 1992; Leekam, 1992). From that age, they will realise that in order to lie successfully they must convince another of the veracity of a false statement (Oldershaw & Bagby, 1997). A girl who has broken a toy may simply accuse her brother of this transgression. However, she may also actually try to lead her mother into believing that her brother broke the toy, for example, by arguing that she is not strong enough to do it. Children have to be somewhat older to use the latter strategy.

Studies examining the individual verbal cues to deception discussed in this chapter hardly exist with children as participants. In one study lying children were less consistent in successive interviews than children who told the truth, whereas no differences emerged in the number of contradictions that occurred in the statements of the truth tellers and liars (Bruck, Ceci, & Hembrooke, 2002). In another study, however, lying children were more consistent than children who told the truth (Quas, Davis, Goodman, & Myers, 2007; Strömwall & Granhag, 2005, 2007; see also Fivush, Peterson, & Schwarzmueller, 2002). In two studies children's lies sounded less plausible than their truths (Ball & O'Callaghan, 2001; Landström, Granhag, & Hartwig, 2007).

[4] See Kendall and Tannen (2001) and O'Connell, Kowall, and Dill (2004) for gender differences in speech.

Regarding personality, Chapter 2 revealed that people high in Machiavellianism lie frequently in order to achieve their goals. One might therefore expect people high in Machiavellianism to have superior verbal lying skills. There is, however, no evidence for this. For example, truth tellers and liars high in Machiavellianism showed similar differences regarding negative statements, self-references, other-references, response length, and lexical diversity as truth tellers and liars low in Machiavellianism (Knapp, Hart, & Dennis, 1974; O'Hair, Cody, & McLaughlin, 1983). No differences were found concerning other personality characteristics either. Riggio and Friedman (1983) found no differences between introverts and extraverts, or between good and poor actors, in the extent to which providing plausible answers was a diagnostic cue to deceit. DePaulo, Epstein, and LeMay (1990) examined differences in response length between truth tellers and liars who scored low or high in social anxiety. Similar differences in response length between truth tellers and liars emerged in both social anxiety groups.

I think it is too premature to conclude that no personality related differences exist in individual verbal cues to deception. For example, perhaps intelligence plays a role. In most studies the participants were university students whose intelligence, we may assume, will be above average. It will be interesting to see whether individual verbal cues are more strongly related to deception in less intelligent people. Theoretically, this may be possible. Less intelligent people may have more difficulty in fabricating credible answers, and so individual verbal cues that are associated with cognitive load (generalising terms, self-references, immediacy, response length, plausibility, contradictions, or consistencies) may therefore be more prominent in liars with a lower IQ.

CONCLUSION

I started this chapter by mentioning that speech content can be revealing about deception, particularly when the liar says something that the observer knows is not true. A straightforward comparison between the facts and the statement of the alleged liar reveals deceit beyond doubt in such a situation.

This chapter further shows that, even when factual evidence is absent, speech content can still provide information about deception. This is intriguing because typically not much is said about the relationship between deception and speech content in police manuals, and the relationship between nonverbal behaviour and deception is discussed in more detail in such manuals. In part this may be because, particularly in the 1970s and early 1980s, researchers used to pay considerably less attention to verbal cues to deception than to nonverbal cues to deception. However, this has changed since the introduction of Criteria Based Content Analysis in the late 1980s, and Reality Monitoring in the 1990s, two verbal assessment tools that I will describe in Chapters 8 and 9.

I believe that police manuals, and people in general, tend to neglect paying attention to verbal cues to deception, because it is assumed that liars are able to control their speech well and therefore are unlikely to leak deception through their speech. I think it is incorrect to assume that liars always control their speech well. First, although people will be aware of what they are conveying, they may be less aware of their exact wording. As a result, they may not notice minor changes in their speech when they start lying. People are unlikely to attempt to control their speech if they don't notice changes in their speech. Second, the task of fabricating a convincing lie may be too difficult, and, as a result, verbal cues to deceit may occur. Even eloquent speakers may reveal verbal cues to deceit due to cognitive load. For example, Clinton was asked in the Paula Jones case whether anyone else other than his attorneys told him that Monica Lewinsky had been served with a subpoena in that case. His answer under oath was "I don't think so". This answer is in conflict with Vernon Jordan's testimony (Clinton's friend), as Jordan had said that he discussed Lewinsky's subpoena with the President. When confronted with this contradiction, Clinton gave, in my view, an implausible answer by saying that he was trying to remember who the *first* person was who told him about Lewinsky's subpoena. This answer is unconvincing because in that case one would expect him to give a name of a person and not to say "I don't think so."

Appendix 4.1 Individual verbal cues to deception

	Negative statements	Generalising terms	Self-references	Immediacy	Response length	Plausible answers	Lexical diversity	Consistencies	Contradictions
Anolli & Ciceri (1997)					>				
Anolli *et al.* (2003)					>				
Bond *et al.* (1985)	–		–						
Bond *et al.* (2005)	–		–						
Burgoon *et al.* (1996a)				<					
Burgoon *et al.* (1996b, study 1)			<	<					
Burgoon *et al.* (1996b, study 2)			<	<					
Burgoon & Qin (2006, 1st part of interview)			–		>		<		
Burgoon & Qin (2006, 2nd part of interview)			–		<		<		
Caso *et al.* (2006b)					>				
Cody & O'Hair (1983)					<				
Cody *et al.* (1984, immediate response)		>			<		–		
Cody *et al.* (1984, delayed response)		–			<		–		
Cody *et al.* (1989)		>							
Colwell *et al.* (2002)					<		>		
Davis *et al.* (2005)					–				
DePaulo P. J. & DePaulo (1989)	–	–							
DePaulo, Epstein *et al.* (1990)					<				
DePaulo, LeMay *et al.* (1991)						–			
DePaulo, Rosenthal *et al.* (1982b)	>	–	<						
deTurck & Miller (1985)					<				
Driscoll (1994)			–						

Dulaney (1982)	–		–		<		>		
Ebesu & Miller (1994)		–	<	<	<				
Feeley & deTurck (1998)			–		<				
Granhag & Strömwall (2002, within one person comparison)					<			–	–
Granhag *et al.* (2003, two suspects comparison)								>	–
Granhag *et al.* (2003, within one person comparison)								–	–
Greene *et al.* (1985)					<				
Harrison *et al.* (1978)					>				
Heilveil & Muehleman (1981)					<				
Hershkowitz (1991)					–				
Klaver *et al.* (2007)					–				
Knapp *et al.* (1974)	>	>	–		<		<		
Kraut (1978)					<	<			
Kraut & Poe (1980)				–					
Landström *et al.* (2005)						–			
Matarazzo *et al.* (1970)					–				
Mehrabian (1971, exp. 1)					–				
Mehrabian (1971, exp. 2)					<				
Mehrabian (1971, exp. 3)					<				
Miller *et al.* (1983)					–				
Motley (1974)					<				
Newman *et al.* (2003)	>		<						
O'Hair *et al.* (1981)					<				
Porter & Yuille (1996)			–		–		–		
Porter, Doucette *et al.* (in press)			–						
Rassin & van der Sleen (2005, within one person comparison)								<	
Riggio & Friedman (1983)						<			
Roberts *et al.* (1998)			<						
Rockwell *et al.* (1997)					<				
Ruby & Brigham (1998)					–				

(Continued)

Appendix 4.1 *(Continued)*

	Negative statements	Generalising terms	Self-references	Immediacy	Response length	Plausible answers	Lexical diversity	Consistencies	Contradictions
Santtila *et al.* (2000)					<				
Schooler *et al.* (1986)			<		<				
Sporer (1997)					–				
Sporer & Sharman (2006)					>				
Stiff & Miller (1986)			–		–	<			
Strömwall, Bengtsson *et al.* (2004)					<				
Vrij (2006)					–				
Vrij, Edward *et al.* (2000)					<				
Vrij *et al.* (2004a)					–				
Vrij *et al.* (2006a)			–	>					
Vrij, Mann *et al.* (2007)					–				
Wagenaar & Dalderop (1994, two suspects comparison)								>	
Winkel & Vrij (1995)					<				
Zhou *et al.* (2004a)	–	–	–		>		<		
Zhou *et al.* (2004b)	>				<				
Zhou & Zang (2006)					<				
Zuckerman *et al.* (1979)			<						
Summary	>	–	–	<	<	<	–	–	–
DePaulo, Lindsay *et al.* (2003)	**.21**	.10	–.03	**–.55**	–.03	**–.23**	–.10		

< cue is shown less by liars than by truth tellers
> cue is shown more by liars than by truth tellers
– no differences was found between liars and truth tellers
Empty cells mean that the cue was not investigated
DePaulo, Lindsay *et al.* (2003) did not investigate consistencies or contradictions

CHAPTER 5

Beliefs About Nonverbal and Verbal Cues to Deception

> A liar rubs his big toe on the ground, and shivers; his face is discoloured; he rubs the roots of his hairs with his fingers; and he tries by every means to leave the house (The Vedas books of Hindu holy knowledge, 900 BC, Seager & Wiseman, 1999, p. 33).

In Chapters 3 and 4 I discussed how liars behave and what they say. Those chapters revealed that typical deceptive nonverbal and verbal behaviour does not exist, although certain nonverbal and verbal cues are weakly related to deception. The present chapter addresses people's beliefs about how liars behave and what they say. Do people think that liars behave nervously or do they think that liars stay calm? And do people think that liars' statements are consistent or do they believe that liars are likely to contradict themselves? Such questions will be covered in this chapter. Measuring people's beliefs about cues to deception is important. The beliefs that people hold often guide their behavioural intentions (Eichenbaum & Bodkin, 2000; Strömwall, Granhag, & Hartwig, 2004). Measuring beliefs about cues to deception might therefore predict which cues people use when they attempt to detect deceit, and, consequently, might predict how good a person might be at accurately detecting lies. Indeed, Forrest, Feldman, and Tyler (2004) found that people with more accurate beliefs about cues to deception made better lie detectors than people with less accurate beliefs.

I start this chapter with a summary of the research literature examining the beliefs people have about nonverbal and verbal cues to deception. In this section I also consider how widely dispersed such views are. To what extent do people from different countries share the same beliefs? And to what extent do people with different professional backgrounds hold the same beliefs? This section reveals that the views held across the world are remarkably similar. Also, professional lie catchers such as police officers, customs officers, immigration officers, and prison officers hold similar views, and their views do not differ from those held by non-professionals such as university students. This section finally demonstrates that there is one group whose views differ somewhat from the view of others, and these are prisoners.

A belief could be defined as a strong or weak feeling or conviction that something is true or real (Strömwall, Granhag, & Hartwig, 2004). This definition clearly implies that beliefs are not necessarily accurate. In the second section of this chapter I compare the beliefs people have about nonverbal and verbal cues to deception with the findings about how liars actually behave and the things they tend to actually say. I will show that both laypersons' and professional lie catchers' beliefs about the nonverbal and verbal behaviour of liars tends to be quite different from the reality. This raises the questions about from where these incorrect beliefs originate and why they seem to last. I answer these questions in the final sections of this chapter. I will argue that many beliefs originate from the notion that "lying is bad" and give reasons as to why incorrect beliefs are likely to persist once they are established.

PEOPLE'S BELIEFS ABOUT DECEPTIVE BEHAVIOUR AND SPEECH

Three Ways of Measuring Beliefs

Beliefs about cues to deception can be measured in different ways. For example, people could be asked to answer the open-ended question "How can you tell when someone is lying?" This *open-ended question method* is straightforward but has not been often used, though was employed in a unique study measuring people's beliefs worldwide, and also in two studies in which one measured the beliefs of British police officers and the other of German police detectives.[1]

[1] A slightly different version of this open-ended question method is used as part of lie detection experiments. In such experiments, participants are shown video fragments of people who are telling the truth or lying, and have to judge after each fragment whether they thought it was a truth or a lie. Participants are then asked which cues they used to

A limitation of asking an open-ended question is that people may have beliefs about cues to deception that do not occur to them when they are asked such a question. For this reason, many researchers ask participants to complete questionnaires that contain closed questions. In a typical *closed question method* questionnaire, cues are listed and participants are asked to indicate their belief about the relationship between the cue (e.g., "eye contact") and deception using rating scales for their answers with headings such as: "much more when lying", "more when lying", "slightly more when lying", "no difference", "slightly more when telling the truth", "more when telling the truth", and "much more when telling the truth".

In another method to measure beliefs about cues to deception, *the correlation method*, researchers compile videotapes of truth tellers and liars and code in detail the nonverbal and verbal behaviour displayed by these truth tellers and liars. Those tapes are then shown to participants (lie detectors) who are asked to judge after each fragment whether they believed the person was telling the truth or lying. The judgements of the lie detectors are then correlated with the actual nonverbal and verbal cues that were present in each video fragment. The outcomes tell us which cues people actually use to indicate whether someone is lying. For example, when there is a tendency amongst lie detectors to judge those who moved a great deal as more deceptive than those who made few movements, we can conclude that they used making movements as a cue to deceit.

One could question whether this method actually measures beliefs about deception. It is unclear whether lie detectors know which cues they actually use, that is, in the example above, whether they were actually aware that they associated making movements with deceit. It is also unknown whether they are guided by their beliefs when they judge the videotapes. That is, in the example above, someone may believe that liars typically decrease their movements, but may, for whatever reason, decide that in the experiment those persons who made many movements were lying. However, since people's behaviour is generally guided by their beliefs, this method should provide insight into people's beliefs. An overlap in results between the first two methods (open-ended

come to their decision. In this version of the open-ended question method the answers given are not context-free but depend on the video fragments that the observers had to judge. For example, it is unlikely that observers will report any verbal cues if the truth tellers and liars provided short answers. Compared to the open-ended question version described in the main text, this version is therefore more restricted. Also, when we compared these two versions in our study, we found an overlap in the results (Mann, Vrij, & Bull, 2004). I will therefore not discuss this restricted version of the open-ended question method in this chapter.

question and closed question methods) and this correlation method is therefore expected.

The Open-Ended Question Method

Charles Bond conducted an ambitious "beliefs about cues to deception" project that he unselfishly published under the name "The Global Deception Team". He recruited an international team of researchers from 58 countries. Each researcher collected data from 20 male and 20 female residents of their country. All participants were over 16 years old and most of them were university students.

The participants were asked to write down their response to the question: "How can you tell when people are lying?" They mentioned 103 different beliefs, and nine of those were given by more than 15% of the participants. Table 5.1, first column, reveals that one cue in particular was prevalent: gaze aversion. People overwhelmingly believed that liars avert their gaze, and this belief was expressed by 64% of the participants. There was considerable worldwide consensus about this belief. Gaze aversion was the most frequently mentioned belief in 51 out of 58 countries. It showed the lowest prevalence in the United Arab Emirates, where even so it was mentioned by 20% of the participants, making it the eighth strongest belief in that country.

Other beliefs mentioned by more than 25% of the participants were that liars are nervous, that their statements are incoherent, and that

Table 5.1 Beliefs about cues to deception: open-ended question

The Global Deception Team (2006)	Mann *et al.* (2004)	Greuel (1992)
Mentioned by more 25% of more of the participants		
Gaze aversion (64%)	Gaze aversion (73%)	Inconsistencies/lack of plausibility (87%)
Nervous (28%)	Body movements (25%)	Atypical victim behaviour (48%)
Incoherent (25%)		
Body movements (25%)		
Mentioned by between 15 and 25% of the participants		
Facial expression	Self-adaptors	
Inconsistencies	Vagueness	
Hesitations	Inconsistencies	
Facial colouration	Hesitations/pauses	
Pauses	Voice (pitch/volume)	
	Face (sweating/blushing/ blinking)	

they make many body movements. Moreover, lies were thought to be given away via facial expressions, verbal inconsistencies, speech hesitations, facial colouration, and pauses. There was also considerable consensus amongst the participants of all 58 countries about those cues and it can thus be concluded that people hold such beliefs worldwide. Those beliefs are predominantly signs of nervousness (gaze aversion, nervousness, incoherent answers, body movements, facial expressions, speech hesitations, facial colouration), and thus suggest that people believe that liars will act more nervously than truth tellers.[2]

In another study, we interviewed 99 British police officers who were on average 34 years old and had served an average 11 years in the police (Mann, Vrij, & Bull, 2004). We asked them to write down their answer to the open-ended question: "What verbal or nonverbal cues do you use to decide whether another person is lying or telling the truth?" A total of 30 different beliefs emerged, and eight of them were mentioned by more than 15% of the police officers. The second column of Table 5.1 shows that gaze aversion was the most popular belief. A majority of 73% of the police officers mentioned that liars avert their gaze. The second most mentioned cue was making body movements, and this was mentioned by 25% of participants. Between 15% and 25% of the police officers further thought that liars make self-adaptors, provide vague and inconsistent statements, hesitate or pause, show changes in voice, and sweat, blush or blink.

The overlap between this survey and the survey conducted by The Global Deception Team is considerable. For example, gaze aversion and making body movements emerged amongst the most prominent cues in both surveys, and also inconsistencies, hesitations and pauses were amongst the most mentioned cues in both surveys. This gives the impression that professional lie catchers (participants in Mann *et al.*'s survey) have similar beliefs about cues to deception as laypersons (participants in the worldwide study). Both professionals and laypersons appear to predominantly believe that liars act more nervously than truth tellers.

Greuel (1992) carried out a more specific "beliefs about cues to deception" study. She asked 51 German police detectives, who all specialised in the investigation of sexual assault, the question: "On which cues do you base your judgement concerning the truthfulness of rape

[2] Some of those cues (gaze aversion, incoherent answers, verbal inconsistencies, speech hesitations, and pauses) could also be seen as signs of cognitive load. However, the results do not show a consistent pattern regarding cognitive load. For example, thinking hard is associated with a *de*crease in movements, but people typically think that liars *in*crease their movements. My preferred interpretation of the data is therefore that people predominantly believe that liars show nervous behaviours.

complaints?" Popular answers were "availability of other evidence" (mentioned by 48% of the respondents) or "intuition" (mentioned by 31% of the respondents). Regarding beliefs about nonverbal and verbal cues to deception two responses emerged. The vast majority of respondents (87%) reported that inconsistencies in statements or a lack of plausibility of statements raised their suspicions (Table 5.1, third column). Many officers further mentioned that "atypical behaviour" displayed by victims evoked their suspicion.[3]

The Closed Question Method

To my knowledge, 16 closed questions surveys regarding beliefs about cues to deception have been published in English to date (the 15 studies reported in Appendix 5.1 and Taylor & Hill-Davies, 2004)[4]. Those surveys have been conducted in the Netherlands, Sweden, the United Kingdom, the United States, and across the world (The Global Deception Team survey). A variety of groups participated in those surveys. The majority of surveys included laypersons, of whom most were university students. Many surveys also included different groups of professionals such as (in alphabetical order): customs officers, judges, managers, migration board personnel, police officers, prison officers, prosecutors, social workers, and teachers. Finally, in some of the surveys prisoners also participated.

Sometimes the beliefs of different groups of professionals and laypersons were investigated within the same survey. None of these studies found consistent differences between different groups of professionals; neither did the beliefs of the professionals differ from those of laypersons. However, a different picture emerged for prisoners and their beliefs differed somewhat from the beliefs of both professional lie catchers and laypersons. For that reason, when reporting beliefs, I differentiate only between prisoners and all the other groups combined.

In their survey, Taylor and Hick (2007) distinguished between "trivial" and "serious" lies. Some differences emerged in participants' beliefs regarding cues to deception in trivial and serious lies, and I therefore report the results for these two types of lies separately. All studies examined the beliefs about cues to deception in adults (as did

[3]There is more evidence that professional lie catchers look for atypical victim behaviour to determine the truthfulness of victims' statements. This strategy is part of Statement Validity Assessment, a verbal assessment tool that I discuss in Chapter 8. As I report in that chapter, this strategy is worrying, because looking for atypical victim behaviour implies that typical truthful victim behaviour exists. This is not the case.

[4]The Taylor and Hill-Davies (2004) findings were not reported in a manner that would enable me to present them in the format used in Appendix 5.1.

the open-ended question studies). In one study, however, the beliefs about cues to deception in young children (five to six year olds) and adolescents (14 to 15 year olds) were examined in addition to those in adults (Vrij, Akehurst, & Knight, 2006). People's beliefs do not appear to differ for varying age groups (see also Taylor & Hill-Davies, 2004), and the results for the different age groups are therefore combined.

A summary of the closed questions surveys is presented in Appendix 5.1. A "<" sign means that observers associate a decrease of the particular cue with deception; a ">" sign indicates that observers associate an increase of the particular cue with deception; and a "−" sign means that participants do not believe that the particular cue is related to deception. Empty cells mean that the cue was not investigated in that study.

Appendix 5.1 shows that participants in different countries share the same beliefs about cues to deception. They believe that compared to truth tellers, liars make more speech disturbances (speech hesitations and speech errors), have a higher pitched voice, a faster speech rate, a longer latency period, and more pauses. They further believe that liars are less able to maintain eye contact, and make more movements (self adaptors, illustrators, movements of the hands, fingers, feet, legs, trunk, head, shifting positions, and eye blinks). People also believe that liars' statements are less immediate, plausible, or consistent, and contain more contradictions. Most of these suggested cues are related to nervousness (more speech hesitations and speech errors, higher pitched voice, faster speech rate, less eye contact, more movements, and less immediate answers), therefore suggesting that people predominantly think that liars act more nervously than truth tellers.[5]

Non-prisoners endorse these views more than prisoners do (Granhag, Andersson, Strömwall, & Hartwig, 2004; Vrij & Semin, 1996, see Appendix 5.1). The differences between non-prisoners and prisoners are larger than Appendix 5.1 suggests. For example, although both non-prisoners and prisoners believe that liars avert their gaze, non-prisoners endorse this view more than prisoners (Vrij & Semin, 1996). Finally, Taylor and Hick's (2007) findings suggest that people expect more signs of nervousness when liars tell serious lies than when they tell trivial lies.

[5]Again, some of those cues (more speech hesitations and speech errors, longer latency period, more pauses, gaze aversion, implausible and less immediate answers, more inconsistencies, and more contradictions) could also be seen as signs of cognitive load. But again, the results do not show a consistent pattern regarding cognitive load (a *slower* speech rate and a *de*crease of movements could be expected when cognitive load is experienced rather than *faster* speech and an *in*crease of movements) and so my preferred interpretation of the data is that people predominantly believe that liars show nervous behaviours.

The Correlation Method

Like the survey studies, the correlation studies have been conducted in different countries, including Germany, the Netherlands, Sweden, the United Kingdom, and the United States. In most studies the participants were university students, but in other studies they were customs officers (Kraut & Poe, 1980), parole officers (Ruback, 1981; Ruback & Hopper, 1986), or police officers (Vrij, 1993; Vrij, Akehurst, Van Dalen, Van Wijngaarden, & Foppes, 1996; Vrij, Edward, & Bull, 2001b; Vrij, Foppes, Volger, & Winkel, 1992a,b, 1994; Vrij, Winkel, & Koppelaar, 1991). As Appendix 5.2 shows, once again, no systematic differences emerged between different countries and different groups of respondents.

Appendix 5.2 shows that observers associate an increase in speech hesitations and speech errors, a high-pitched voice, a slower speech rate, longer pauses, an increase in gaze aversion, self-adaptors, hand and finger movements, leg and foot movements, trunk movements, and shifting positions, and fewer head movements with deceit. Observers further associate fewer self-references, less immediate, shorter and less plausible answers, and more inconsistencies with deception.

Many cues that emerge in the correlation studies (Appendix 5.2) show overlap with the cues that people believe to be signs of deception (closed question method, Appendix 5.1).[6] The overlap in findings between Appendix 5.1 and Appendix 5.2 supports the idea that observers' beliefs about cues to deception (Appendix 5.1) guide the cues they focus on when attempting to detect deceit (Appendix 5.2).

For other cues (speech rate, latency period, pause durations, illustrators, head movements, and response length) differences emerged between the closed question and correlation methods. Those beliefs may thus be less clear-cut than the closed question method suggests.[7]

Many findings of the correlation studies (liars show more speech hesitations and speech errors, a higher pitched voice, more gaze aversion, more movements, fewer self-references, less immediate and shorter answers) again support the notion that people predominantly believe that liars act more nervously than truth tellers. However, several correlation studies indicate that people also associate "odd behaviour" with deception, that is, behaviours that are deemed deviant or inappropriate.

[6] Speech hesitations, speech errors, pitch of voice, gaze, self-adaptors, hand and finger movements, leg and foot movements, trunk movements, shifting positions, immediacy of speech, plausibility, and consistencies.

[7] Some other cues (pause frequency, eye blinks, negative statements, generalising items, self-references, lexical diversity, and contradictions) have not been measured enough to make a comparison between both methods possible.

For example, three studies have shown that not only excessive gaze aversion, but also staring makes a suspicious impression (Bond, Omar, Pitre, Lashley, Skaggs, & Kirk, 1992; Desforges & Lee, 1995; Levine *et al.*, 2000). Both total gaze aversion and staring violate norms and fall outside the range of behaviours typically considered as normal. In a similar vein, Baskett and Freedle (1974) and Boltz (2005) found that responses that came too slowly or too quickly were perceived as deceptive, whereas the responses that followed after an intermediate delay were thought to be truthful. Finally, Kraut (1978) found that both an excessive amount of hesitations and an absence of hesitations raised suspicion.

In summary, a suspect in a police interview, a husband who is challenged by his wife about an affair, or a child who denies wrongdoing to her parents, will make a suspicious impression when displaying signs of nervousness. However, they will also raise suspicion if they show odd behaviour such as too much eye contact, answering questions too quickly, or speaking too smoothly (see also Henningsen, Cruz, & Morr, 2000).[8]

ACCURACY OF BELIEFS

To what extent are the beliefs people have about cues to deception accurate? In this section I compare the beliefs people have about cues to deception with liars' actual behaviour (Chapter 3) and speech (Chapter 4). I have summarised the findings of Chapters 3 and 4 in the second column of Table 5.2 and labelled them "objective indicators of deception". In the third column of Table 5.2 I have summarised the findings regarding beliefs about cues to deception, discussed in this chapter. I labelled these beliefs "subjective indicators of deception".

[8]The situation is somewhat more complicated than I just presented in the text because the context in which the behaviour takes place has an effect on how behaviour is interpreted. That is, some behaviours look more suspicious in one context than in another (Aune, Levine, Ching, & Yoshimoto, 1993; Feldman & Chesley, 1984; Kraut, 1980; Kurasawa, 1988; Strömwall & Granhag, 2003a). In one experiment, observers saw one of two simulated video-dating service interviews with a female. In the video the general appearance of the female was manipulated so that her desirability as a partner was either emphasised or minimised. That is, her appearance was either typical of a young, fashion-conscious woman on a dinner date or inappropriate for a dating situation. In both videos, the female displayed cues stereotypically associated with deception, such as looking away from the interviewer, and shifting her posture (the same behaviour was showed in both interviews). In both interviews the female mentioned that she is an adventurous person who really enjoyed blind dating. Observers rated the female as more deceptive when her appearance was inappropriate for a dining situation (Aune *et al.*, 1993).

Table 5.2 Objective and subjective indicators of deception

	Objective (actual) indicators[a]	Subjective indicators (beliefs)[b]
Vocal cues		
Hesitations	–	>
Speech errors	–	>
High pitched voice	>	>
Speech rate	–	–
Latency period	>	–
Pause durations	>	–
Pause frequency	–	>
Visual cues		
Gaze	–	<
Smiling	–	–
Self-adaptors	–	>
Illustrators	<	–
Hand/finger	<	>
Leg/foot	<	>
Trunk	–	>
Head	–	–
Position shifts	–	>
Eye blinks	–	>
Verbal cues		
Negative statements	>	–
Self-references	–	<
Immediacy	<	<
Response length	<	–
Plausible answers	<	<
Consistencies	–	<
Contradictions	–	>

[a] Explanation of the signs: < shown less often by liars than by truth tellers; > shown more often by liars than by truth tellers; - no relationship with deception
[b] Explanation of the signs: < observers believe that liars show the cue less often than truth tellers; > observers believe that liars show the cue more often than truth tellers; - observers do not associate the cue with deception.

In the second column, a "<" sign indicates that a particular cue is shown less by liars than by truth tellers; a ">" sign means that a particular cue is shown more by liars than by truth tellers; and a "–" sign means that no difference was found between liars and truth tellers. In the third column a "<" sign indicates that observers believe that liars show the cue less than truth tellers; a ">" sign means that observers believe that liars show the cue more than truth tellers; and a "–" sign means that observers do not associate the cue with deception.[9]

Table 5.2 shows that people's views are often incorrect. Out of the 24 cues reported in Table 5.2, people hold correct beliefs about only six cues (pitch of voice, speech rate, smiles,[10] head movements, immediacy, and plausibility). People thus hold correct views about only 25% of the cues listed in Table 5.2.[11] Table 5.2 further reveals that people overestimate the number of cues that are related to deception. Whereas only 10 cues show any actual (weak) relationship with deception, people believe that 16 out of the 24 cues are related to deception.

People have often incorrect views about the 10 diagnostic cues to deception. Out of those 10 cues, people are only aware about the relationship with deception for three of them (liars tend to have a higher pitch of voice, and are more inclined to give non-immediate and implausible answers). People therefore have correct beliefs about only 33% of the diagnostic cues listed in Table 5.2. Two cues that are related to deception, hand and finger movements and leg and foot movements, are actually related in a way that opposes people's beliefs: liars tend to exhibit less of these movements whereas people believe they increase such movements. The remaining five cues that are related to deception (latency period, pause durations, illustrators, negative statements, and response length) are not believed to be associated with deception. The latter finding is interesting. People believe that only eight out of the 24 cues listed in Table 5.2 are not associated with deception, yet five of these eight cues are actually related to deception.

In summary, people typically have incorrect beliefs about cues to deception. They associate lying with many cues that have actually no relationship with deception, and of those cues that are to some extent related, people are often unaware of their true relationship with

[9] I did not include the cues "generalising statements" and "lexical diversity" in Table 5.2 because people's beliefs about these cues have not been investigated enough.

[10] The results for the category "smiles overall" is included in Table 5.2, and this category is not related to deception. However, when a distinction is made between felt and false smiles, relationships with deception do occur (Chapter 3). It is unknown whether observers are aware of these relationships.

[11] A distinction between vocal, visual, and verbal cues suggests that people have similar knowledge about vocal cues (29% accuracy), visual cues (20% accuracy), and verbal cues to deception (29% accuracy). I already mentioned that prisoners' beliefs appear to differ from the beliefs of professional lie catchers and undergraduate students. Our own study showed that the prisoners' beliefs are more accurate (Vrij & Semin, 1996). We calculated accuracy scores regarding beliefs about cues to deception held by customs officers, police officers, prison officers, prisoners, and university students. The prisoners held more accurate beliefs than any of the other groups, whereas the accuracy of beliefs of the other groups did not differ from each other.

deception, and often think that they are not indicative of lying. As a result, on the one hand people associate several cues with deception that are in fact unrelated to deception, but, on the other hand, are unaware of several cues that are to some extent related to deception.

One of the main reasons for the mismatch between objective and subjective indicators of deception is that people's views about cues to deception are too simplistic. People predominantly believe that liars will be more nervous than truth tellers and therefore will show more signs of nervousness. In previous chapters I have already discussed the fact that liars are not necessarily more nervous than truth tellers, and furthermore if a liar is nervous, he or she will not necessarily show more signs of nervousness. For example, liars may successfully control their behaviour and speech and, as a result, do not show such signs. Moreover, signs of nervousness may be absent because the cognitive load associated with telling the lie may automatically suppresses the occurrence of such signs.

Box 5.1 More subjective cues to deception

Research has revealed more subjective cues to deception than those presented in the main text of this chapter. Particularly the extensive closed question surveys of Akehurst, Köhnken, Vrij, & Bull (1996), Lakhani and Taylor (2003), Taylor and Hick (2007) and Vrij, Akehurst, and Knight (2006) shed light on such cues. The pattern found in these surveys about the additional cues is identical to what has been discussed so far, in that people associate many more signs of nervousness with deception, whereas most of them are not actually associated with deception. Those who are interested in these cues should read these surveys.

Other studies revealed that clothing can affect suspicion. I showed 91 Dutch police detectives, of whom 92% reported having considerable experience in interviewing suspects, video fragments of a series of truth tellers and liars (Vrij, 1993). After each fragment I asked the detectives to indicate whether the person was telling the truth or lying. The results showed that the detectives' judgements were influenced by the way the mock suspects in the video fragments were dressed. The suspects who were untidily dressed made a more suspicious impression than those who were smartly dressed. In another study we found that a person wearing black clothing made a more suspicious impression on observers than a person who wore light clothing (Vrij & Akehurst, 1997). This finding adds to the research literature demonstrating that wearing black clothing makes a negative impression on others (Frank & Gilovich, 1998; Vrij, 1997; Vrij, Pannell, & Ost, 2005). Needless to say, the way people are dressed is not a reliable indicator of deception.

Yet another set of deception studies revealed that facial appearance impacts on the impression that someone makes. People with attractive faces are typically thought of as more honest (Aune, Levine, Ching, & Yoshimoto, 1993; Bull & Rumsey, 1988), as are individuals with a baby-faced appearance (high foreheads and widely-spaced eyes, Masip, Garrido, & Herrero, 2003a,b, 2004; Zebrowitz, Voinescu, & Collins, 1996). These characteristics are not valid cues to deception either.

THE ORIGIN OF INCORRECT BELIEFS ABOUT NONVERBAL AND VERBAL CUES TO DECEPTION

One main finding of this chapter is that people hold many incorrect beliefs about nonverbal and verbal cues to deception. This raises the question of where these cues originate. I believe there are three sources, which I will call the moral, exposure, and accusation explanations.

The Moral Explanation

As I discussed in Chapter 2, the stereotypical view is that lying is bad. Charles Bond argues that the belief that liars avoid eye contact fits well with this lying-is-bad stereotype (The Global Deception Team, 2006). If lying is bad, then people should feel ashamed when they lie. People often avert their gaze when they feel ashamed (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). Moreover, if lying is bad, then people should feel nervous of getting caught when they lie, resulting in signs of nervousness such as avoiding eye contact, but also in additional nervous cues such as stuttering, talking with a higher pitched voice, and moving about more.

This moral reasoning could explain why children as young as five to six years old already associate gaze aversion and limb movements with deception (Rotenberg & Sullivan, 2003). From a young age, children are taught that it is wrong to lie, and, subsequently, children will associate cues of shame and anxiety with lying. Moral reasoning can also explain why prisoners endorse less than others the belief that liars act nervously. Criminals probably need to lie frequently to succeed in their criminal careers, and therefore probably feel less bad about lying than non-criminals. If they do not feel bad about lying, there is less reason to expect liars to act nervously.

The Exposure Explanation

The idea that liars look away or show other behaviours associated with nervousness is prominent in the popular media. Virtually all articles that I read about deception in the popular media express this view in some shape or form. Thus, in case people have not come to the conclusion themselves that liars act more nervously than truth tellers, those articles will put this notion in their head.

It is not just the popular media that promotes this idea. Police manuals also do. Pär Anders Granhag and myself reviewed a large number of police interrogation manuals (Vrij & Granhag, 2007). Although such manuals typically contain small warnings about the unreliability of cues to deception, these are easily lost in the subsequent detailed and enthusiastic descriptions of how speech and behaviour differs between truth tellers and liars (see also Moston, 1992). Table 5.3 gives an overview of the nonverbal cues these manuals suggest looking for.

Table 5.3 Examples of nonverbal cues to deception found in published police interrogation manuals

Manual	Nonverbal cues to deception
Gordon & Fleisher (2002)	Problem with eye contact Touching the nose Restless foot and leg movements
Inbau, Reid, Buckley, & Jayne (2001)	Avoiding eye contact Frequent posture changes Grooming gestures Placing hands over mouth/eyes
Macdonald & Michaud (1992)	Rubbing the eyes Avoiding eye contact Covering/rubbing the ears
Rabon (1992)	Restless behaviour Tapping of feet Fidgeting Excessive swallowing Avoiding direct gaze
Yeschke (1997)	Shuffling the feet Avoiding eye contact Picking lint from clothes High frequency of blinking
Zulawski & Wicklander (1993)	Moving the chair Abrupt and jerky behaviour Problem with fine motor coordination Cold and clammy hands Using hands to cover mouth Failure to maintain eye contact

Source: derived from Vrij and Granhag (2007)

The cues mentioned in these police manuals show great overlap regarding gaze aversion. Every single manual mentions that liars avert their gaze. As I discussed in Chapter 3, there is no evidence for this, including in police-suspect interview situations. Some of the remaining cues are also mentioned in more than one manual (e.g., using hands to cover the mouth). There is no evidence that liars do this either. The cues mentioned in these manuals show overlap because they all support the notion that liars will behave nervously (see also Blair & Kooi, 2004). However, different manuals emphasise different cues of nervousness. It would be interesting to know where these differences originate. It is clear that they are not derived from the deception literature, because not a single cue mentioned in Table 5.3 has emerged as a diagnostic cue to deceit in such literature. Sometimes the amount of detail expressed in these manuals is striking. For example, Gordon and Fleisher (2002) distinguish between women and gay men on the one hand and heterosexual men on the other hand, and claim that women and gay men make different types of self-adaptors to heterosexual men. However, they do not support their claims with any data (Strömwall, Granhag, & Hartwig, 2004).

The Accusation Explanation

I finally believe that the origin of the stereotypical view that liars behave more nervously than truth tellers could arise from the act of accusing someone of lying. Suppose that a police detective is convinced that a suspect is lying, but the suspect keeps on denying any wrongdoing during the police interview. In order to break the suspect's resistance the police detective may decide to accuse the suspect of lying. As Chapter 3 revealed, accusations alone could easily result in the suspect showing nervous behaviours such as gaze aversion and an increase in movements, and this is likely to occur in both innocent and guilty suspects (Bond & Fahey, 1987; Ekman, 1985/2001). The police detective, however, may not realise that his or her accusation will make the suspect behave nervously and may assume that the suspect shows such behaviour as the result of lying.

WHY INCORRECT BELIEFS ABOUT NONVERBAL AND VERBAL CUES TO DECEPTION LAST

Once incorrect beliefs have been established they are difficult to discard. As a result, they will endure. This happens for a variety of reasons, and I will discuss five of them.

Illusory Correlations

Once incorrect views have been established, people will perceive supporting evidence that in fact does not exist. For example, once observers have formed the impression that someone is lying, they then overestimate the amount of gaze aversion the alleged liar actually displays (Levine, Asada, & Park, 2006). Thus, they think that the alleged liar is showing gaze aversion (e.g., perceiving supporting evidence) even when this in fact is not the case.

Our experiments provide another example of perceiving supporting, but non-existing, evidence. In these experiments, where participants told the truth and lied on camera, we asked participants how they believe they behaved when they were telling the truth and lying, and then examined their actual behaviour ourselves (Vrij, Semin, & Bull, 1996; Vrij, Edward, & Bull, 2001a). Although the participants showed fewer movements when they lied compared to when they told the truth, they thought that they had moved more whilst lying. Also, whereas the participants showed similar gaze patterns during truth telling and lying, they thought they had averted their gaze more when they were lying. In other words, people believe that they themselves show cues of nervousness when they lie, such as increased movements and gaze aversion (e.g., perceiving supporting evidence) when they in fact do not exhibit such behaviours.

Apart from perceiving supporting evidence that does not exist, illusory correlations may explain in another way why the incorrect belief that liars lack eye contact is so persistent. Gaze patterns are highly noticeable. Conversation rules dictate that we look people in the eye when we talk to them, which makes it likely that we pay attention to gaze patterns and will try to find out whether there is a relationship between gaze patterns and deception. This pattern is in fact erratic. Some suspects display less eye contact when they are lying whereas others display more eye contact when they are lying. Different patterns of gaze aversion during truth telling and lying also emerge within individuals, and the convicted murderer showed a clear example of this (Chapter 3). He looked away a lot while lying before his confession, and maintained eye contact while lying during his confession.

Despite that there is no relationship between gaze behaviour and lying, people will not easily admit this because they have the tendency to seek explanations for, and to create order and predictability in, ambiguous events (Gilovich, 1991). This may eventually lead to perceiving relationships that do not actually exist. People tend to overestimate the frequency of occurrence of events that are meaningful to them

(Chapman, 1967). Since the idea that liars look away is meaningful to people, they will overestimate the number of times they noticed that liars looked away. They then eventually will see a relationship between gaze aversion and deception that does not actually exist.

Confirmation Bias

People tend to seek information that confirms rather than disconfirms their beliefs (so-called *confirmation bias*, Darley & Gross, 1983). Any support they find for their beliefs will boost their confidence that their views are correct, making it less likely that they will alter them. Customs officers who believe that smugglers tend to look away and move a lot, will search the luggage of passengers who show this behaviour rather than searching the luggage of people who look them into the eyes and stay calm. Since some smugglers will lack eye contact and display an excessive amount of movements, the officers are likely to find illegal products in the luggage of at least some of the passengers they search. This will boost their confidence that they use the correct strategy in whose luggage to search, and they will be inclined to continue using this strategy. When visiting one country I spoke to a senior officer of the military police, an organisation responsible for checking passports at airports. He told me that he thought that his organisation was good at seizing false passports, and to substantiate his opinion he showed me a videotape of passengers who were showing their passports to a military police officer. One woman was clearly nervous which raised the officer's suspicion. He carefully checked her passport and discovered that it was a false document. The senior officer thought that this example demonstrated that the looking-for-cues-of-nervousness strategy works.

In fact, it does not demonstrate this. I asked him whether one or more of the other passengers on the tape who did not show nervous behaviour were also in the possession of a false document. The senior officer replied that he did not know because the passports of these passengers had not been thoroughly examined. This is a shame, because only by checking the passports of a great number of people who do show nervous behaviour and also checking the passports of a great number of people who do not show nervous behaviour it is possible to assess whether the looking-for-cues-of-nervousness strategy works. For this strategy to be effective it must be demonstrated that those who stay calm do not possess false passports. However, checking the passports of people who stay calm is seeking information that disconfirms most people's beliefs, and hence is unlikely to happen.

Belief Perseverance

When people come across an example that disconfirms their beliefs, they are more likely to disregard it than interpret this new evidence as a sign that their initial belief is incorrect. This phenomenon is called *belief perseverance* (Anderson, Lepper, & Ross, 1980), and is another reason why people's established incorrect beliefs are likely to persist. People may be confronted with a liar who maintained eye contact and did not show nervous behaviours, but they are unlikely to give this observation the importance it deserves. They may perceive this as an exceptional case or may have another reason to explain away this evidence.[12]

The Power of Thinking

Once people have formed an opinion that makes sense to them, they will come up with further reasons to support their view. If people are asked why they think liars look away, they may think of reasons to corroborate this view and search their memory for examples where they encountered liars who averted their gaze (Strömwall, Granhag, & Hartwig, 2004). Thinking about examples that support their beliefs will strengthen their opinion that liars look away. In other words, an opinion is often strengthened by merely thinking about the topic (Tesser, 1978).

Poor Feedback

I will discuss one more reason why incorrect established views often last. People often do not receive the adequate feedback they need to learn from their own experience and to discover that their views are inaccurate. In order for feedback to be effective, it needs to be given frequently, reliably, and immediately. In terms of feedback about nonverbal and verbal cues to deception, observers should be informed immediately after every interaction with another person whether that person was lying or not so that they could work out how liars truly behave and what they really say. However, this is not a realistic option, as people do not discover whether or not they have been lied to in most interactions. If

[12]The confirmation bias and belief perseverance are related to the *need for cognitive closure* (Ask, 2006; Ask & Granhag, 2005, 2007; Kruglanski & Webster, 1991). Need for cognitive closure (NFC) refers to the desire for a clear-cut opinion on a topic. Both situational and individual factors influence NFC. Regarding situational factors, NFC is heightened when the task is boring, or when the observer is under time pressure or tired (Ask & Granhag, 2007). Regarding individual factors, individuals with high NFC tend to make quick decisions because they dislike the uncertainty that precedes the decision. They also tend to stick by their decisions to prevent that uncertainty from reoccurring (Ask & Granhag, 2007).

they do come to discover that they have been lied to, it is often a long time after the interaction took place (Park, Levine, McCornack, Morrisson, & Ferrara, 2002), by which time they have probably forgotten how the person behaved and what the person said exactly. Interestingly, those who search luggage and check passports are in a position to give themselves adequate feedback, because they could immediately find out whether someone is lying. To create this adequate feedback they need to search the luggage and check the passports of people randomly, including those who they do not believe to be smuggling or holding false passports. As I explained above, people are unlikely to do this.

CONCLUSION

People from different countries and with different professional backgrounds share beliefs about how liars behave and what they say. Above all, they believe that liars are unable to maintain eye contact. More generally, they predominantly believe that liars act more nervously than truth tellers. When I compared those beliefs with how liars actually behave and what they actually say, it became evident that people's beliefs are often incorrect. People associate more cues with deception than is justified. Also, for those cues that are to some extent indicative of lying, people are often unaware of how they actually relate to deception, or consider them not to be related to deception at all. As a result, people associate several cues with lying that are in fact unrelated, but are unaware of several other cues that are to some extent diagnostic cues to deceit.

I further gave reasons where these incorrect beliefs originate from and one explanation was that it is caused by the stereotypical view that lying is bad. If lying is bad, then people should feel bad when they lie, and should feel nervous about getting caught. I then discussed reasons why incorrect beliefs about deception are likely to last once they have been established: illusory correlations (perceiving relationships that actually do not exist); confirmation bias (a tendency to seek information that confirms existing beliefs); beliefs perseverance (a tendency to disregard evidence that opposes existing beliefs); the power of thinking (thinking about evidence and reasons that support someone's beliefs); and poor feedback (inadequate feedback that hinders people from learning from their own experience) all make it likely that incorrect beliefs will endure.

I believe that the combination of how incorrect beliefs originate and why they last could explain why incorrect beliefs about cues to deception appear in many police manuals. I have no doubt that the authors of these manuals believe that the information they provide (i.e., liars are nervous and will show nervous behaviours) is correct. However, their views are based on their own or other police officers' impressions about how suspects behave and what they say during police interviews rather than on systematic research. These impressions can easily become distorted as this chapter revealed. My advice to authors of police manuals therefore is to base their writing on systematic research rather than on impressions.

Measuring beliefs about cues to deception is important because, as this chapter also demonstrated, it provides insight into the strategies that lie detectors use when they attempt to detect deceit. In that respect, the findings are not encouraging and suggest that lie detectors often use ineffective strategies when paying attention to nonverbal and verbal cues, which may result in poor performances in lie detection. Chapter 6 demonstrates that this pessimistic view is often a reality.

Appendix 5.1 Subjective nonverbal cues to deception: closed questions

	Vocal Cues						
	Hesitations	Speech errors	High-pitch voice	Speech rate	Latency period	Pause durations	Pause frequency
Closed Questions							
Akehurst *et al.* (1996, 'other' ratings)	>	>	>	<	>		>
Colwell *et al.* (2006)	>	>	>	>	>		>
Gordon *et al.* (1987)	>	>		>			
Granhag *et al.* (2004, prisoners)			–			–	
Granhag *et al.* (2004, non-prisoners)			>			<	
Granhag *et al.* (2005)						<	
Lakhani & Taylor (2003, high stakes)	>	>	>				>
Strömwall & Granhag (2003b)			–				
Taylor & Hick (2007, trivial)	–	<	–				<
Taylor & Hick (2007, serious)	>	–	>				>
Taylor & Vrij (2001)	>	>	>	>	>		
The Global Deception Team (2006)		>			>		
Vrij, Akehurst, & Knight (2006)	>	>	>	>	>		>
Vrij & Semin (1996, prisoners)	>	>	–	>	–		
Vrij & Semin (1996, non-prisoners)	>	>	>	>	–		
Zuckerman *et al.* (1981b, 'other' ratings)	>	>	>	>	>		
Summary	>	>	>	>	>	–	>

Appendix 5.1 (*Continued*)

	Visual Cues									
	Gaze	Smile	Self adaptors	Illustrators	Hand/ finger	Leg/ foot	Trunk	Head	Shifting position	Eye blink
Closed Questions										
Akehurst *et al.* (1999, 'other' ratings)	<	>	>	>		>	>	>	>	>
Colwell *et al.* (2006)	<	–	>	>	>	>	>		>	>
Gordon *et al.* (1987)	<	>	>	>		–	>			>
Granhag *et al.* (2004, prisoners)	<									
Granhag *et al.* (2004, non-prisoners)	<									
Granhag *et al.* (2005)	–									
Hart *et al.* (2006)	<	–		>	>	>		>	>	>
Lakhani & Taylor (2003, high stakes)	<	<	>		>	>		>	>	>
Sromwall & Granhag (2003b)	<									
Taylor & Hick (2007, trivial)	–	–	<		–	–		–	–	–
Taylor & Hick (2007, serious)	>	–	–		>	–		–	–	–
Taylor & Vrij (2001)	<	–	>	>	>	>	>	>	>	>
The Global Deception Team (2006)	<		>	>					>	
Vrij, Akehurst, & Knight (2006)	<	<	>	–	>	>	>	>	>	>
Vrij & Semin (1996, prisoners)	<	–	>	–	–	>	–	>	–	
Vrij & Semin (1996, non-prisoners)	<	>	>	>	>	>	>	>	>	
Vrij & Taylor (2003)	–	–								
Zuckerman *et al.* (1981b, 'other' ratings)	<	>	>	>		>		>	>	>
Summary	<	–	>	>	>	>	>	>	>	>

Appendix 5.1 (*Continued*)

	Individual Verbal Cues								
	Negative statements	Generalising terms	Self-references	Immediacy	Response length	Plausible answers	Lexical diversity	Consistencies	Contradictions
Closed Questions									
Akehurst *et al.* (1996)				<	<	–			>
Colwell *et al.* (2006)				<	>				
Granhag *et al.* (2004, prisoners)								–	
Granhag *et al.* (2004, non-prisoners)								<	
Granhag *et al.* (2005)								<	
Lakhani & Taylor (2003, high stakes)					>			–	>
Stromwall & Granhag (2003b)								<	
The Global Deception Team (2006)					>			<	
Vrij, Akehurst, & Knight (2006)	>			<	–	<			>
Vrij & Taylor (2003)						<			
Zuckerman *et al.* (1981b, 'other' rating)	–		>	<					
Summary				<	–	<		<	>

< observers associate a decrease of the cue with deception
> observers associate an increase of the cue with deception
– observers do not believe that the cue is related to deception
Empty cells mean that the cue was not investigated

Appendix 5.2 Subjective nonverbal cues to deception: correlation studies

	Vocal Cues						
	Hesitations	Speech errors	High-pitch voice	Speech rate	Latency period	Pause durations	Pause frequency
Correlation Studies							
Apple *et al.* (1979)			>	<			
Basket & Freedle (1974)		>			>,<		
Boltz (2005)					>,<		
Bond *et al.* (1985)	–	–					
Bond *et al.* (1990)	–						>
Bond *et al.* (2004)	>			<			
DePaulo, Rosenthal *et al.* (1982)	>	>		<			
DePaulo, P. J., & DePaulo (1989)	>	–		<			
Ekman (1988)			–				
Fiedler & Walka (1993)	–		>	<			
Frank & Ekman (2004)				<			
Harrison *et al.* (1978)					<		
Kraut (1978, exp. 1)	–	–			>		
Kraut (1978, exp 2)	>,<						
Kraut & Poe (1980)		–			>		
McCroskey & Mehrley (1969)		–					
Nigro *et al.* (1989)	>	>					
Riggio & Friedman (1983)	–	–		<			
Riggio *et al.* (1987)				<			
Ruback & Hopper (1986)						>	
Ruva & Bryant (1998)	>						
Stiff & Miller (1986)	>				–	>	
Streeter *et al.* (1977)			>				
Vrij (1993)	–	–	–				
Vrij, Foppes *et al.* (1992)	>	>					
Vrij & Winkel (1994)	>	>					
Vrij *et al.* (2001b)	>	–		–	>		
Woodall & Burgoon (1983)				<			
Summary	>	>	>	<	–	>	

Appendix 5.2 (*Continued*)

	Visual Cues									
	Gaze	Smile	Self adaptors	Illustrators	Hand/ finger	Leg/ foot	Trunk	Head	Shifting position	Eye blink
Correlation Studies										
Akehurst & Vrij (1999)					>					
Bond *et al.* (1985)	<	>	–	–				–	–	
Bond *et al.* (1990)	<	–	>	<				<		–
Bond *et al.* (1992)	>,<									
Brooks *et al.* (1986)	–									
DePaulo, P. J., & DePaulo (1989)	–		–					–	>	
Desforges & Lee (1995)	>,<									
Ekman & Friesen (1972)			>							
Ekman (1988)		–	–	–		–				
Fiedler & Walka (1993)		–	–					–		
Frank & Ekman (2004)			>	<						
Hemsley & Doob (1978)	<									
Hess & Kleck (1994)	<									>
Kraut (1978, exp 1)		>	<				>		>	
Kraut & Poe (1980)	<	–	–	–			>		>	
O'Sullivan *et al.* (1988)				>		>				
Riggio & Friedman (1983)	–	<	>	–		–		<	–	
Riggio *et al.* (1987)	<	–						<		
Rozelle & Baxter (1978)	<			>			>			
Ruback (1981)		<		<						
Ruback & Hopper (1986)			<	–						
Stiff & Miller (1986)		>	–	–		–			>	–
Vrij (1993)	–	<			>	–	–	–		
Vrij, Akehurst *et al.* (1996)			>		>	>				
Vrij, Foppes *et al.* (1992)	<	>	–		–			–	–	
Vrij, Winkel *et al.* (1991)	<									
Vrij & Winkel (1992)		>	>	>						
Vrij *et al.* (2001b)	–	–	–	–	–	>				
Summary	<	–	>	–	>	>	>	<	>	

Appendix 5.2 (*Continued*)

	Individual Verbal Cues								
	Negative statements	Generalising terms	Self-references	Immediacy	Response length	Plausible answers	Lexical diversity	Consistencies	Contradictions
Correlation Studies									
Bond *et al.* (1985)	–		<						
Bond *et al.* (1990)	<				–				
Bond *et al.* (2004)				<					
DePaulo, Rosenthal *et al.* (1982)	>	>	<						
DePaulo, P. J., & DePaulo (1989)	–	–		–					
Granhag & Strömwall (2000a)								<	
Harrison *et al.* (1978)					>				
Kraut (1978, exp 1)						<			
Kraut & Poe (1980)				<					
Leippe *et al.* (1992)								<	
Lind *et al.* (1978)					<				
Riggio *et al.* (1987)			<		<				
Stiff & Miller (1986)			–		<	<			
Stiff *et al.* (1989)						<			
Vrij (1993)					–				
Summary	–		<	<	<	<		<	

< observers associate a decrease of the cue with deception
> observers associate an increase of the cue with deception
– observers do not believe that the cue is related to deception
Empty cells mean that the cue was not investigated

CHAPTER 6

Lie Detection Without Using Specialised Tools

This is the first of eight chapters examining people's ability to detect lies. In Chapters 7 to 13 I discuss the accuracy of specialised tools that are used by professional lie catchers and scholars. The current chapter examines people's ability to catch liars without the use of such specialised tools, and when they pay attention to someone's nonverbal and verbal behaviour.

When I discuss people's ability to detect lies, I refer to their ability to *discriminate* between truths and lies. A good lie detector is not only able to judge when someone is lying, but also when someone is telling the truth. Suppose that someone has to assess the veracity of 10 statements of which five are actually truthful and five are deceptive. If he were to judge all 10 statements as deceptive then he would correctly identify all five lies, but does this make him a good lie detector? Not according to my definition, because he made incorrect judgements about all five truthful statements. I would be impressed by his skills, however, if he correctly identified all five lies *and* all five truths.

In this chapter I discuss how good laypersons are at detecting truths and lies in: (i) people they do not know; (ii) their own friends and partners; and (iii) children. We will see that laypersons' performance is only marginally better than what can be expected by chance, that is, what could be expected by flipping a coin. Although laypersons can to some extent successfully indicate when someone is telling the truth, they are poor at identifying liars, and this is the case in evaluating strangers, friends, partners, and children. I will then turn to professional lie catchers. This chapter reveals that they too perform at a level that is only slightly better than could be expected by chance when they identify truths and lies told by people unknown to them. They are somewhat better than laypersons at judging that someone is lying, but at the expense of correctly identifying truth tellers. Perhaps the most striking differences between professional lie catchers and laypersons are that professionals are: (i) more suspicious than laypersons (more inclined to call someone a liar) and (ii) more confident in their ability to discriminate between truths ands lies.

Despite this pessimistic view about people's ability to detect truths and lies, studies have revealed individual differences in this ability. I discuss factors that could possibly explain these differences, and they include lie detector's gender, personality, and motivation.

I further discuss several factors that influence the veracity judgements observers make (i.e., do they believe that is someone is telling the truth or lying). Non-Caucasian people often make a suspicious impression on Caucasian observers, due to the behaviour non-Caucasian people naturally display. In addition, people with certain personality traits tend to be believed more than people lacking those traits. Again, the natural behaviour shown by these individuals causes this effect. I also demonstrate that: (i) actually interviewing someone rather than passively observing that person and (ii) further questioning ("probing") that person influences veracity judgements. Interviewers are more credulous than passive observers and probing further increases the tendency to believe someone.

I commence this chapter with a discussion about the relative impact of nonverbal and verbal communication on veracity assessments. That is, if people try to detect deceit, what influences their decision the most: the behaviour shown by the target person or the content of his or her speech? I will demonstrate that this question cannot be answered in general terms.

THE RELATIVE IMPORTANCE OF NONVERBAL AND VERBAL BEHAVIOUR IN CREDIBILITY ASSESSMENTS

Several sources (police manuals, research findings, and real-life observations) indicate that nonverbal behaviour plays an important role in making veracity judgements. Police manuals typically allocate more pages to nonverbal cues than to verbal cues to deceit. In fact, verbal veracity tools such as Criteria-Based Content Analysis (Chapter 8) and Reality Monitoring (Chapter 9), which have been shown to discriminate to some extent between liars and truth tellers, are not mentioned in such manuals. The importance of nonverbal cues is further emphasised in police manuals with statements such as "as much as 70 percent of a message communicated between persons occurs at the nonverbal level" (Inbau, Reid, Buckley, & Jayne, 2001, p. 143).

Regarding research findings, in our lie detection experiment we showed 99 British police officers 54 videotaped fragments of police interviews with suspects. We asked them to make veracity judgements after each fragment and to report the cues on which they based their decisions (Mann, Vrij, & Bull, 2004). The vast majority of the cues the police officers reported (78%) were nonverbal. Also, when observers notice that someone's nonverbal behaviour is discrepant from his or her speech content they typically rely on the nonverbal channel. Thus, if in a selection interview an applicant with a reserved demeanour claims to be enthusiastic about the job, people tend to believe that the applicant is not as keen as he or she claims (DePaulo, Rosenthal, Eisenstat, Rogers, & Finkelstein, 1978; Hale & Stiff, 1990; Heinrich & Borkenau, 1998; Hoffner, Cantor, & Thorson, 1989; Zuckerman, Driver, & Koestner, 1982; Zuckerman, Speigel, DePaulo, Rosenthal, 1982).

Several real-life observations further demonstrate the importance of nonverbal behaviour in judging whether someone is lying. Kaufmann, Drevland, Wessel, Overskeid, and Magnussen (2003, p. 22) noticed that in their home country, Norway, judicial decisions are sometimes based on nonverbal communication, even when it contradicts available evidence. They describe a court trial in which the circumstantial evidence of guilt was strong, but where the defendant (a financial adviser) was acquitted partly because "his nonverbal behaviour was confident without evasive eye movements of any sort". In Florida, Tom Sawyer was accused of sexual assault and murder. Investigators interrogated him for 16 hours, issued threats, and extracted a, probably false, confession. He became a prime suspect because he appeared embarrassed and his face flushed during an initial interview in which he denied involvement in the crime (Meissner & Kassin, 2002). (See Ofshe, 1989, and Box 11.2 for a more detailed description of the Tom Sawyer case.) In another American case, 14-year-old Michael Crowe was submitted to lengthy interrogations and confessed to having stabbed his sister to death. The charge against the boy was later dropped. He became a prime suspect because the detectives believed that he had reacted to his sister's death with inappropriately little emotion (Kassin, 2005).

Why do lie detectors pay so much attention to nonverbal behaviour? First, they may feel confident in their ability to interpret it. By observing behaviour alone, people can determine with reasonable accuracy all sorts of things about other people, including personality traits (e.g., extraversion and sociability), masculinity, femininity, or sexual orientation. From behaviour it is also possible to discern information about status, dominance, and romantic involvement (Ambady, Bernieri, & Richeson, 2000; Ambady & Rosenthal, 1992; DePaulo, 1992; DePaulo & Friedman, 1998). In other words, people are used to making inferences from nonverbal behaviour.

Second, perhaps observers pay so much attention to nonverbal behaviour because they assume that people are less aware of it than their speech and that, as a result, they "leak" the information they try to hide through nonverbal channels (DePaulo *et al.*, 1978; Hale & Stiff, 1990; Kalbfleisch, 1992; Maxwell, Cook, & Burr, 1985; Stiff, Hale, Garlick, & Rogan, 1990; Vrij, Dragt, & Koppelaar, 1992). Indeed, under certain circumstances individuals are less able to control some aspects of their nonverbal behaviour than their verbal communication (DePaulo, 1992; DePaulo & Kirkendol, 1989; Ekman, 1993; Ekman & Friesen, 1969, 1974). Suppose that at the airport a customs officer asks a heroin smuggler whether he has anything illegal in his suitcase. It is rather easy for the smuggler to say that he has nothing to hide, but it is probably more difficult for him to behave normally and avoid raising any suspicion through his behaviour. Nor will it be difficult for the student to tell the invigilator that her lecture notes that he found on the floor during the exam do not belong to her, but it will be more difficult for her to stay calm. I already discussed in Chapter 3 why it is difficult for liars to control their behaviour in these circumstances. Amongst other difficulties, they need to suppress signs of nervousness whilst showing credible behaviour.

Predominantly paying attention to nonverbal behaviour, however, is not what lie detectors always do. Research has shown that people sometimes pay more attention to nonverbal behaviour and at other times to speech content when making veracity judgements, and the relative importance of the two channels depends on many factors. For example, it depends on what the observer knows about the topic of conversation (Stiff, Miller, Sleight, Mongeau, Garlick, & Rogan, 1989). In cases where the observer is knowledgeable, he or she typically concentrates on the speech and compares what he or she knows with what the target person says. In Chapter 4 I gave an example of how Jeffrey Archer's lie about his alleged telephone call with the Prime Minister was caught in this way.

It also depends on the amount of speech content available to observers. Sometimes a person will only say a few words or a couple of sentences, as in the examples of the heroin smuggler and student above. In those cases an observer has almost no other choice than to examine someone's behaviour. In other situations there is more speech content available to observers. For example, observers may have access to more than one statement. They may have obtained more than one statement about an alleged event from the same person, or they may have obtained statements about an alleged event from different persons. In such cases, lie detectors tend to focus on speech content, checking for consistency between the different statements (Granhag & Strömwall, 1999, 2000a,b, 2001a,b; Strömwall & Granhag, 2005, 2007; Strömwall, Granhag, & Jonsson, 2003).

Another factor influencing the relative impact of nonverbal and verbal cues in making veracity judgements is the distinctiveness of these cues. For example, people rely on the verbal channel when they find a story implausible (Kraut, 1980), or when a statement appears to be against the self-interest of the story teller (Noller, 1985). In contrast, when someone displays behaviour that is perceived as odd, such as staring, their veracity judgements are likely to be influenced by these odd behaviours (Bond, Omar, Pitre, Lashley, Skaggs, & Kirk, 1992).

Moreover, expectancies about the truthfulness of a person may influence what people pay attention to. Police officers readily assume that a suspect is guilty (Evans, 1994; Kassin, 2005; Moston, Stephenson, & Williamson, 1992; Stephenson & Moston, 1994). Analyses of police interviews in England showed that in 73% of the cases the police interviewers were "certain" of the suspect's guilt before interviewing them (Moston *et al.*, 1992); and Saul Kassin (2005, p. 216), who over the years has asked numerous American police officers whether they are concerned that their persuasive interrogation methods may make innocent people confess, reported that the most common reply he received is: "No, because I do not interrogate innocent people". When lying is expected, police officers may have little interest in listening to a suspect's flat denials and prefer to look at bodily signs to detect deceit (Millar & Millar, 1998).

Finally,[1] different people may evaluate the verbal and nonverbal channels differently (Friedman, 1978). Noller (1985) reported that females are less responsive to speech content than men and rely more

[1] I am aware that I could discuss more factors. For example, whether the lie detector is an actual interviewer or a passive observer might be another variable influencing whether to pay attention to speech content or nonverbal communication. The findings, however, are inconsistent. Buller, Strzyzewski, and Hunsaker (1991) and Feeley and deTurck (1997) found that active interviewers pay most attention to nonverbal communication, whereas observers mostly listen to speech content. Granhag and Strömwall (2001b) found the opposite effect.

on nonverbal cues. Women are typically more knowledgeable than men about nonverbal behaviour and better at interpreting it (Hall, 1984). This enhanced ability may guide their attention to nonverbal cues. In a similar way, lie detectors' knowledge about verbal and nonverbal cues to deceit may affect their focus of attention. Those who perceive themselves as knowledgeable about verbal cues to deception may listen carefully to what someone says, whereas those who think they know more about nonverbal cues to deception may carefully scrutinise someone's behaviour.

As I mentioned above, police manuals emphasise nonverbal cues to deception more than verbal cues. This could mean that police officers tend to pay more attention to nonverbal cues to deception, even in situations where listening to the speech content would be more appropriate. When they start their investigations, police officers sometimes have little information about the crime (Horvath & Meesig, 1996), and therefore do not yet have a clear idea about who the suspects are. Therefore, they may decide to interview a number of people who inhabit the area where the crime took place, or a number of people who they believe could have committed the crime. The nonverbal style of presentation of these individuals is likely to determine whether they will go on to be considered as prime suspects and invited for a second interview (Greuel, 1992; Kraut & Poe, 1980; Rozelle & Baxter, 1975; Vrij, Foppes, Volger, & Winkel, 1992; Walkley, 1985; Waltman, 1983). This is what happened in the cases of Tom Sawyer and Michael Crowe as described earlier. I believe this approach has limitations. As I discussed in Chapter 4, mental states, such as negative feelings, can also leak through via speech, and producing plausible answers on the spot is sometimes difficult. Controlling speech may also be difficult and lie detectors' strategy to predominantly pay attention to nonverbal behaviour may thus be too restrictive in such circumstances. Also, as this chapter reveals, people are often better at differentiating between truths and lies when they listen to, rather than watch, someone.

In summary, whether the observer pays attention to the verbal or nonverbal channel depends on what the observer knows about the topic of conversation, the amount of speech content available to observers, the distinctiveness of the verbal and nonverbal cues, the observer's expectations about whether the person is lying, and the observer's knowledge about verbal and nonverbal cues of deceit.

LAYPERSONS' ABILITY TO DETECT LIES IN STRANGERS

Laypersons' ability to detect lies in strangers has been examined extensively in the last few decades. In a typical lie detection study, observers (normally undergraduate students) are given short video fragments of people they do not know who are either telling the truth or lying. They are asked to indicate after each fragment whether the person (often called the *sender*) was telling the truth or lying. Typically, half of the senders are truth tellers, and half are liars.[2] In such a study, simply guessing whether the sender spoke the truth or lied would result in correctly classifying 50% of the truths (truth accuracy rate) and 50% of the lies (lie accuracy rate), resulting in a total accuracy rate of 50%. The suspicious observer introduced in the introduction paragraph of this chapter who judges all 10 statements as deceptive where in fact five were deceptive and five were truthful, would obtain a lie accuracy rate of 100%, a truth accuracy rate of 0% and a total accuracy rate of 50%. The latter score indicates that his performance was identical to the level of chance (50%).

In lie detection studies, observers are not given any background information about the senders and their statements, so the only source of information available to them is the nonverbal and verbal behaviour displayed by these senders. Compared to real life this is a somewhat unusual way of detecting lies. In their study, Park, Levine, McCornack, Morrisson, & Ferrara (2002) asked college students to recall an instance in their life in which they had detected that another person had lied to them, and to report how they had discovered the lie. Less than 2% of the lies were detected at the time the lie was told by relying exclusively on the liars' nonverbal behaviour or speech content. More commonly, lies were discovered via information from third parties (38%), physical evidence (23%), and confessions (14%). More than 80% of the lies were detected one hour or more after they were told, and 40% were detected more than a week later.

In many lie detection studies, the fragments the observers have to judge were derived from the studies that have been discussed in Chapters 3 and 4. In most of these studies senders told the truth or lied for the sake of the experiment. Liars typically experienced some cognitive load and the stakes for truth tellers and liars varied from minor to moderate.

Kraut (1980) published a review of studies concerning lie detection. The total accuracy rates in most of these studies ranged from 45% to 60%, and the average accuracy rate was 57%. Many studies have been published since Kraut's review appeared in 1980. Appendix 6.1 presents the 79 studies that I am aware of where laypersons' ability to discriminate between truths and lies told by people they did not know, and

[2]The observers are typically not informed what percentage will be truth tellers and liars, because this may result in them deliberately trying to achieve an equal number of truth/lie responses.

which have been published in English since 1980. The lowest total accuracy rate was 31% (Brandt, Miller, & Hocking, 1982), and this appears to be an outlier. The second lowest total accuracy rate was 38%, and was obtained in another study by Brandt, Miller, and Hocking (1982). The highest total accuracy rate ever reported was 68% (Wan Cheng & Broadhurst, 2005). The remaining total accuracy rates ranged from 42% to 65%. Appendix 6.1 further shows that in the vast majority of studies (62 out of 79) total accuracy rates between 50% and 60% were reported. The average accuracy rate of these 79 studies was 54.27% and this is somewhat lower than the 57% accuracy rate found by Kraut (1980). An accuracy rate of 54.27% is only just above the 50% chance level.

Appendix 6.1 further shows a distinction has been made between truth accuracy (correct classifications of truths) and lie accuracy (correct classifications of lies) in 34 studies. Truth accuracy rates are typically higher than lie accuracy rates: In 25 of these 34 studies the truth accuracy rates were higher than the lie accuracy rates. The truth accuracy rates ranged from 49% to 81%. Truth accuracy rates above 65% are common and were achieved in 16 out of 34 studies. The average truth accuracy rate was 63.41% and this is higher than could be expected by chance. In contrast, the highest lie accuracy reported was 70% but lie accuracy rates as low as 27% have also been obtained. Lie accuracy rates below 50% are common and were achieved in 18 of 34 studies. The average lie accuracy rate was 48.15%, and this is below that that could be expected by chance.

Truth-Bias

The superior truth accuracy rate is at least in part the result of a *truth-bias*, which is the tendency of observers to judge messages as truthful rather than deceptive (Köhnken, 1989; Levine, Park, & McCornack, 1999; Zuckerman, DePaulo, & Rosenthal, 1981). In their meta-analysis of lie detection studies, Bond and DePaulo (2006) found that observers judged 56% of messages as honest and only 44% as deceptive, despite being exposed to an equal number of truths and lies. In other words, observers tend to be credulous. There are at least eight explanations for the truth-bias (DePaulo, Wetzel, Weylin Sternglanz, & Walker Wilson, 2003; Gilbert, 1991; Vrij, 2004b; Vrij & Mann, 2003a). First, in daily life people are more often confronted with truthful than with deceptive statements (Chapter 2), so they are more inclined to assume that someone is telling the truth (the so-called *availability heuristic*, O'Sullivan, Ekman, & Friesen, 1988).[3] Second, as outlined in Chapter 2, Goffman

[3] Heuristics are simple decision rules that permit the evaluation of complex stimuli with limited cognitive effort (Tversky & Kahneman, 1974).

(1959) argued that life is like a theatre and that people often behave as actors and put on a show. The "self" that is presented to others in daily life is not the true self but an edited version. In such a theatre, we expect others to honour the way we present ourselves, but we also accept the way others present themselves to us. The latter point results in a truth-bias.

Third, social conversation rules discourage people from displaying suspicion (Toris & DePaulo, 1984). A person would become quickly irritated if their conversation partner questioned everything being said. Imagine a conversation in which someone interrupts you all the time by saying things like "I don't believe you", "That cannot be true", and "Could you prove that?" The conversation is unlikely to last long. In other words, social conversation rules dictate credulity. Fourth, people may be unsure as to whether deception is actually occurring. Given this uncertainty, the safest and most polite strategy may be to believe what is overtly expressed (DePaulo, Jordan, Irvine, & Laser, 1982). Related to this is Fiedler and Walka's (1993) *falsifiability heuristic*, a fifth explanation of the truth-bias. Fielder and Walka argued that statements that are easily falsifiable via reality checks appear less credible than messages that are not easily falsifiable. People, however, often lie about topics that are not easily falsifiable, such as their feelings, preferences, attitudes, and opinions (Chapter 2). In those cases, observers tend to take the messages for granted.

Sixth, based on a Spinozan model of knowledge representation, Gilbert (1991) argued that everything is initially taken to be true, and that disbelief requires extra effort. In other words, the truth-bias is the default setting in social interactions. Related to this, Elaad (2003) presented a seventh reason for the existence of the truth-bias, based on Tversky and Kahneman's (1974) *anchoring heuristic*. It refers to people making insufficient adjustments from an initial value (the anchor) resulting in a final decision that is biased towards this value. Thus, if observers are preoccupied in thinking that someone is telling the truth, they will make insufficient adjustments when contrasting evidence emerges. Finally, Grice (1989) argued that in social interactions, the language people use is expected to follow a set of principles (see also McCornack, 1992). For example, the quantity principle dictates that information should be enlightening, and the relation principle emphasises that information should be relevant. However, people often violate such rules, and often speak in ambiguous ways. Conversation partners are used to it, allow for it, and often do not become suspicious when this occurs. As a consequence, even if a liar's speech violates language principles then observers will not easily become suspicious.

Although the truth-bias is frequently found, it is not absolute and can be eliminated by raising suspiciousness (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). For example, in a study where observers judged statements of salespersons selling their products, they showed a lie-bias (DePaulo & DePaulo, 1989), and in other studies the number of lie judgements increased when participants were told that the senders "may not have been completely truthful" (McCornack & Levine, 1990a,b; Millar & Millar, 1997b). Finally, as we will see below, police officers, who often assume guilt in the suspects they interrogate, do not show a truth-bias either.

Deception Medium

Lie detectors can judge senders in different ways. For example, in face-to-face conversations they can hear and see the sender, whereas in telephone calls they can only hear the sender. Researchers have examined whether the channel the observers are exposed to has an effect on their lie detection performance.[4] Bond and DePaulo (2006) reviewed this deception medium research and discovered that it does have an effect. Lie detectors who could only hear the senders performed equally well as observers who could both hear and see the senders, but observers who could only see the senders performed worse than observers who could only hear the senders ($d = .37$) or who could both hear and see the senders ($d = .44$) .[5] In other words, observers make worse lie detectors when they cannot hear the senders. There are at least two explanations for this. First, Chapters 3 and 4 revealed more vocal and verbal cues to deception than visual cues, and observers who cannot hear senders do not have access to these more diagnostic vocal and verbal cues. Second, people strongly believe that liars look away and increase their movements (Chapter 5), but these are not diagnostic cues to deceit (Chapter 3). When observers can only see senders they tend to rely on their incorrect stereotypical views when making veracity judgements (Bond & DePaulo, 2006), and incorrect judgements are likely to occur. Also, these stereotypical views are about how *liars* behave rather than how *truth tellers* behave. Therefore, observers who can only see senders are more likely to brand someone a liar than those who can see and hear

[4] See for example Atoum & Al-Simadi, 2000; Bond & Atoum, 2000; Bond, Thomas, & Paulson, 2004; Davis, Markus, & Walters, 2006; DePaulo, Stone, & Lassiter, 1985; Ekman & Friesen, 1974; Ekman, Friesen, O'Sullivan, & Scherer, 1980; Frank & Ekman, 2004; Hocking, Bauchner, Kaminski, & Miller, 1979; Kassin, Meissner, & Norwick, 2005; Maier & Thurber, 1968; Manstead, Wagner, & MacDonald, 1986; Noller, 1985; Porter, Campbell, Stapleton, & Birt, 2002; Zuckerman, Driver, & Guadagno, 1985.

[5] As already mentioned in Chapter 3, effect sizes (d) of .20, .50, and .80 should be interpreted as small, medium, and large effects, respectively.

senders or can only hear senders (Bond & DePaulo, 2006; Mann, Vrij, Fisher, & Robinson, 2007; Vrij, in press).

Motivation/Stakes

I already discussed in Chapter 3 that liars are not all equally motivated to succeed in their deceit. In high-stakes situations where the negative consequences of being disbelieved or the positive consequences of being believed are high, people may be more motivated to appear credible than in low-stakes situations. I also reported in Chapter 3 that being motivated has a detrimental effect on liars: the more motivated liars are to avoid getting caught, the more likely it is that their behaviour will give their lies away, a phenomenon labelled the *motivational impairment effect*. The explanation for this effect is that the three processes that may elicit behavioural responses to deceit (emotions, cognitive effort, and attempted behavioural control) will be more profound in motivated liars than in less motivated liars (Chapter 3).

If liars display more cues to deceit in high-stakes lies than in low-stakes lies, then high-stakes lies should be the easier to detect. Indeed, in a series of experiments where the stakes were manipulated (although the stakes were never really high), observers were better at distinguishing between truths and lies in high-stakes settings than in low-stakes settings.[6] Bond and DePaulo's (2006) review also revealed that lies are somewhat easier to discriminate from truths if they are told by motivated rather than unmotivated senders ($d = .17$).

Preparation

Sometimes people have opportunity to prepare their lie in advance, whereas in other instances they have to lie spontaneously. We already saw in Chapter 3 that planned lies are somewhat easier to tell than spontaneous lies and therefore result in fewer signs of cognitive load. Therefore unsurprisingly, the meta-analysis of Bond and DePaulo (2006) revealed that observers are slightly better at distinguishing between truths and lies when they judge spontaneous messages than when they judge planned messages ($d = .14$).

Familiarity with the Sender

I discussed in Chapter 3 that individual differences exist in how people behave and speak. This makes lie detection in strangers difficult,

[6] DePaulo, Kirkendol, Tang, & O'Brien, 1988; DePaulo, Lanier, & Davis, 1983; DePaulo, LeMay, & Epstein, 1991; DePaulo *et al.*, 1985; Lane & DePaulo, 1999; Vrij, 2000; Vrij, Harden, Terry, Edward, & Bull, 2001.

because observers do not know how these strangers normally behave or speak. Perhaps lie detection is easier if observers come to know the truthful behaviour and speech of the strangers they have to judge so that they can look for changes in behaviour and speech between the message known to be truthful and the message they have to judge. Several studies have been carried out examining this. In those studies observers still had to judge the veracity of statements of senders they did not know, but this time they were shown "truthful baseline messages" of these senders prior to the lie detection task (the observers were informed that the baseline messages were truthful). Those studies indicated that observers who were exposed to truthful baseline interviews performed better at the subsequent lie detection task than observers who were not exposed to such baseline interviews (Brandt, Miller, & Hocking, 1980a, b, 1982; Feeley, deTurck, & Young, 1995). In other words, familiarity with the truthful behaviour of a sender makes it easier to discriminate between truths and lies.

Other researchers showed observers baseline messages that were either truthful or deceptive but did not inform the observers about the truth or lie status of these baseline messages. They found that observing the baseline messages only benefited observers when they were truthful (Garrido & Masip, 2001; O'Sullivan, Ekman, & Friesen (1988). O'Sullivan, Ekman and Friesen (1988) argued that, since observers are inclined to assume that the behaviour they observe is honest (availability heuristic), they are more likely to assume that the first behaviour (i.e., baseline behaviour) they see is honest rather than dishonest. If the first behaviour is honest, then it becomes easier to identify subsequent different behaviour as dishonest. If, however, the first behaviour is actually deceptive but mislabelled as honest, errors may occur when subsequent behaviours are assessed.

Summary

Laypersons only perform slightly better than chance when they attempt to distinguish between truths and lies told by strangers. They also have a tendency to believe those strangers (truth-bias), but the truth-bias disappears when the observers become suspicious. Lie detection accuracy is influenced by several factors. Observers are better at distinguishing truths from lies when: (i) they listen to what senders say rather than how they behave; (ii) the stakes for senders are high rather than low; (iii) the senders' messages are spontaneous rather than planned; and (iv) they are familiar with the truthful baseline behaviour of the senders.

Table 6.1 Accuracy rates of laypersons judging friends or romantic partners[1]

	Accuracy rate (%)		
	Truth	Lie	Total
Anderson, DePaulo, & Ansfield (2002, friends)[2]			54
Bauchner (friends, cited in Miller, Mongeau & Sleight (1986))			74
Fleming & Darley (1991, friends and parents)			49
Levine, Park, & McCornack (1999, not married couples)	82	34	58
McCornack & Levine (1990, not married couples)[3]			53
McCornack & Parks (1986, not married couples)			59
Millar & Millar (1995, Study 1: friends and relatives)[4]			51
Millar & Millar (1995, Study 2: friends and relatives)			53

Notes: [1]Only control conditions are included; [2]Time 1 and time 2 data combined; [3]Low suspicion condition; [4]Voice and video information condition

LAYPERSONS' ABILITY TO DETECT LIES IN FRIENDS AND ROMANTIC PARTNERS

If familiarity with the sender's natural truthful responses facilitates truth and lie detection, then one could expect people to be better at detecting truths and lies in their friends and romantic partners than in strangers. In alignment with this, Boon and McLeod (2001) reported that people believe that they are fairly good at detecting lies in their partners.[7] In several experiments observers were shown video fragments of people they know well, and, like the research with strangers, nonverbal behaviours and speech content were the only sources of information available to them. Table 6.1 reports the findings of the studies I am aware of that have been published in English. These studies do not support the idea that it is easier to discriminate between truths and lies told by friends, family, or lovers than by strangers. Miller, Mongeau, and Sleight (1986) reported a study conducted by Bauchner which showed that friends could detect each other's lies with 74% accuracy. As far as I know, such a high accuracy rate has never been found in research with strangers. However, this finding has not been replicated. In other studies accuracy rates between 49% and 59% were obtained. Such accuracy

[7]Interestingly, people believe that they are more successful in deceiving their partners than their partners are at deceiving them (Boon & McLeod, 2001).

rates are similar to the accuracy rates found when observers attempted to detect truths and lies in strangers. In fact, none of the studies where a direct comparison was made between the ability to detect truths and lies in strangers versus in friends or partners found a difference in accuracy rates (Anderson, DePaulo, & Ansfield, 2002; Buller, Strzyzewski, & Comstock, 1991; Fleming, Darley, Hilton, & Kojetin, 1990; Millar & Millar, 1995). Comadena (1982), who examined the ability to discriminate between truths and lies when judging friends or intimates, found no difference between these two groups either. In other studies, the association between relationship closeness and accuracy was examined. In none of them was an association between closeness and accuracy found (Levine & McCornack, 1992; McCornack & Parks, 1986; Stiff, Kim, & Ramesh, 1992).

Anderson, Ansfield, and DePaulo (1999) gave several reasons why no link between relationship closeness and accuracy at detecting deception seems to exist. They suggested that when close relationship partners attempt to detect deceit in each other, they bring to mind a great deal of information about each other. This information could well be overwhelming and the lie detector may deal with this by processing the information heuristically (i.e., using simple judgemental operations) instead of carefully searching for genuine cues to deceit. Another explanation is that in close relationship interactions the lie detector must simultaneously engage in social cognition (e.g., decoding possible cues to deception) and social behaviour (e.g., appear supportive in those interactions) (Patterson, 1995, 2006). This may be too much for the lie detector to concentrate on and, as a result, valuable cues to deceit may remain unnoticed. A further explanation is that as relationships develop, people become more skilled at crafting communications uniquely designed to fool each other. That is, throughout interactions with their partners, liars have learned to tell a lie in such a way that it is difficult for their partners to detect.

Moreover, a substantial number of lies told in romantic relationships are other-oriented lies told to protect the partner or to make them feel better (Chapter 2). People typically do not feel bad when telling such lies and therefore may not show any cues to deceit when telling them. Also, since other-oriented lies contain information that romantic partners want to hear, they may not be motivated to detect such lies. Finally, people tend to invest emotionally in their relationships, and some information can threaten the relationship. In those instances, people may not see through lies, because they do not want to (I labelled motivational reasons for being unwilling to detect lies *the ostrich effect* in Chapter 1).

The latter two motivational explanations suggest that there may be a truth-bias in most romantic relationships. Indeed, it has been argued that as relationships become more intimate, partners develop a strong tendency to judge the other as truthful, the so-called *relational truth-bias heuristic* (Stiff *et al.*, 1992). McCornack and Parks (1986) and Levine and McCornack (1992) have developed and tested a model to explain this. The core of their argument is that closeness and truth-bias are not directly linked to each other, but that they are both linked with confidence. As soon as the relationship between two people intensifies, they become more confident that they can detect each other's lies ("I know so-and-so very well, I am able to tell whether he or she is lying"). High levels of confidence will then result in the belief that the other person would probably not dare to lie to them ("So-and-so had better watch out, I can detect every lie he or she tells me"). This will result in putting less and less effort into trying to discover whether that person is lying ("I don't have to worry that much, so-and-so does not lie to me anyway") (Stiff *et al.*, 1992).

Cole, Leets, and Bradac (2002) examined whether a relational truth-bias depends on the attachment style of the observer. Attachment style tells us how people perceive romantic relationships (Chapter 2). *Secure*[8] individuals have a positive view of themselves and others. They are comfortable with intimacy and value it in their relationships. *Preoccupied* people tend to hold a more negative view of themselves while idealising their romantic partners. They fall in love easily but typically doubt their partners' long-term commitment and level of interest. They long for closeness in their romantic relationships but are anxious about their partners' willingness to meet their relational needs. *Fearful* individuals tend to hold a negative view of themselves and others. They have difficulty in trusting others, tend to avoid intimacy, and are uncomfortable with closeness in their relationships. Finally, *dismissive* people hold positive views about themselves but negative views about others. They tend to deny the importance of intimacy and strive to achieve autonomy. In summary, both secure and preoccupied individuals desire intimacy from a romantic partner and have a positive view of others. In contrast, fearful and dismissive individuals tend to avoid intimacy and have a negative view of others. Cole *et al.* found that the relational truth-bias was particularly present in observers who have a positive view of others (secure and preoccupied observers).

People's poor ability to detect lies in friends and romantic partners and tendency to believe such friends and partners provides a worrying

[8] The two dimensions avoidance and anxiety described in Chapter 2 result in four attachment styles. *Secure* individuals are low in avoidance and low in anxiety; *preoccupied* people are low in avoidance but high in anxiety; *fearful* individuals are high in avoidance and high in anxiety; and *dismissive* people are high in avoidance but low in anxiety (Griffin & Bartholomew, 1994a, b).

prospect in the fight against terrorism. In the United Kingdom, after the 7 July 2005 bombings in London and failed attempts to create further destruction, people are asked to be vigilant in their own communities, including towards their friends and family members. The findings presented in this section make it doubtful whether people are suited to identify potential terrorists amongst their close friends and family members.

Summary

Laypersons are equally able to detect truths and lies in strangers as they are in their friends and partners. There are several explanations for this including that as relationships develop, people become more skilled at crafting communications uniquely designed to fool each other. Also, as relationship partners become closer, partners become more confident that they will be able to detect the other's lies and this eventually leads them to believe that the other will not lie to them (relational truth-bias heuristic). This tendency to believe the other is stronger in secure and preoccupied individuals than in people with a fearful or dismissive attachment style.

LAYPERSONS' ABILITY TO DETECT LIES IN CHILDREN

When a young child and adult both lie about the same event, which of the two lies is the easiest to detect? Most people will probably say that spotting the young child's lie is easier. Indeed, young children do not yet: (i) have the cognitive abilities to make up a plausible lie; (ii) realise that making a credible impression is important; and (iii) possess the muscular control required to display an honest demeanour (Chapter 3). However, Chapter 3 also suggested that detecting lies in children may not be as straightforward as it sounds. For example, children may not yet realise the negative consequences of being disbelieved and may therefore not experience much fear when they lie. Also, with increasing age, people show more spontaneous emotional facial expressions, which sometimes have to be suppressed in order to conceal deceit. Also, young children typically provide shorter statements than adults when they are invited to recall an experienced event. This may benefit young children when they lie, because the less time liars talk, the less opportunity there is for them to display nonverbal cues to deceit, and the less likely it becomes that they will show such cues (DePaulo, Lindsay *et al.*, 2003). Shorter statements also lead to fewer words and this may decrease the likelihood of verbal cues to deception to arise,

Table 6.2 Accuracy rates of laypersons judging children[1]

	Accuracy rate (%)			
	Age of senders	Truth	Lie	Total
Ball & O'Callaghan (2001, exp 1)	6–12	75	53	60
Ball & O'Callaghan (2001, exp 2 and 3 combined)	6–12			65
Chahal & Cassidy (1995)	8	70	63	66
Edelstein, Luten, Ekman, & Goodman (2005)	5–7			50
Jackson & Granhag (1997)	11–12	72	25	49
Landström, Granhag, & Hartwig (2007)[2]	10–11	53	66	60
Larochette, Chambers, & Craig (2006)	8–12			54
Leach, Talwar, Lee, Bala, & Lindsay (2004, exp 1)	3–11			51
Strömwall & Granhag (2005)	11	77	41	59
Strömwall & Granhag (2007)[3]	12–13	57	68	63
Strömwall, Granhag, & Landström (2007)[4]	11–13	55	48	52
Talwar, Lee, Bala, & Lindsay (2006)	4–7	74	26	50
Vrij, Akehurst, Brown, & Mann (2006)	5–6	61	63	62
Vrij & van Wijngaarden (1994)	5, 9	57	51	54

Notes: [1]Only control conditions are included; [2]Live and video conditions combined; [3]Average across all conditions; [4]Prepared and unprepared statement conditions combined

because words are the carriers of verbal cues to deception (Vrij, Mann, Kristen, & Fisher, 2007).

Several studies have been carried out examining adults' ability to distinguish between truths and lies told by children. Those experiments use the same paradigm as the studies examining the detection of deception in strangers. In other words, adult observers assess the veracity of statements of children they do not know and the only information they can rely upon is how the children behave and what they say. Table 6.2 shows the studies published in English that I am aware of in which observers faced a forced choice test and had to indicate for each child whether they believed he or she was lying or not.[9]

[9]A forced choice test format (e.g., the child is either lying or telling the truth) enables researchers to calculate accuracy rates. In child deception detection research it is common to use other formats than a forced choice test. In many experiments adults had to indicate on Likert scales ranging from (1) definitely not to (7) definitely, the extent to

The total accuracy rates in these studies ranged from 49% to 66%. In most studies truths were easier to detect than lies. The accuracy rates in Table 6.2 are not impressive and give the impression that adults have difficulty in detecting children's truths and lies. In particular, the lie accuracy rate in Jackson and Granhag's (1997) study was low (25%). This may have been caused by the experimental design. Truth tellers saw a film about a group of chimpanzees living in the wild and were asked to recall what they had seen. Liars were asked to pretend that they had seen the film and had to fabricate a story. It appears that the liars did very well, as most of them were judged as truthful. A possible explanation is that most children had seen a film about chimpanzees in the wild before or had seen chimpanzees in the zoo and therefore could fabricate stories which were rich in detail and sounded convincing.

In several studies observers have tried to detect truths and lies in children of different ages, providing an opportunity to test whether children's lies are more difficult to detect as they get older (DePaulo, 1991). The results show general support for this assumption, although gender differences emerged (Feldman, 1979; Feldman, Jenkins, & Popoola, 1979; Feldman, Tomasian, & Coats, 1999; Feldman & White, 1980; Morency & Krauss, 1982; Westcott, Davies, & Clifford, 1991). In their study with 5–12 year olds, Feldman and White (1980) found that in girls, but not in boys, the face revealed less about deception with increasing age, making older girls the best liars. Girls are generally more rewarded for being expressive than boys and thus may have greater opportunity to practise nonverbal displays of emotion. Shennum and Bugental (1982) examined 6–12 year olds and found that with increasing age boys became better than girls in their ability to neutralise negative affect (pretend to be neutral about something they disliked). Boys are taught not to show negative emotions, and therefore may be particularly well trained in neutralising negative emotions.

Those findings may suggest that lie detection is easier in children than in adults, but this appears not to be the case. In two studies adults were requested to detect truths and lies in young children (aged five to seven) and in adults. No differences in accuracy emerged between the two age groups (Edelstein, Luten, Ekman, & Goodman, 2006; Vrij,

which they felt the children were lying. The results of these studies cannot be translated into accuracy scores. However, these results show the same pattern as the outcomes of the forced choice studies described in the main text (Allen & Atkinson, 1978; Bugental, Shennum, Frank, & Ekman., 2001; Feldman, 1979; Feldman, Devin-Sheehan, & Allen, 1978; Feldman, Jenkins, & Popoola, 1979; Feldman & White, 1980; Feldman, White, & Lobato, 1982; Goodman, Batterman-Faunce, Schaaf, & Kenney, 2002; Goodman, Myers, Qin, Quas, Castelli, Redlich, & Rogers, 2006; Leach, Talwar, Lee, Bala, & Lindsay, 2004; Lewis, 1993; Lewis, Stanger, & Sullivan, 1989; Morency & Krauss, 1982; Orcutt, Goodman, Tobey, Batterman-Faunce, & Thomas, 2001; Talwar & Lee, 2002).

Akehurst, Brown, & Mann, 2006). Moreover, the accuracy rates presented in Table 6.2 (ability to detect deceit in children) are comparable to the accuracy rates presented in Appendix 6.1 (ability to detect deceit in adults). These findings suggest that detecting truths and lies in children is equally difficult as detecting truths and lies in adults.

Several researchers examined whether observers who have children are better at detecting children's truths and lies than observers who do not have children. The findings are mixed. Chahal and Cassidy (1995) found that parents were better than non-parents in truth and lie detection in children, but several other studies did not reveal differences (Ball & O'Callaghan, 2001; Talwar & Lee, 2002; Vrij, Akehurst, Brown, & Mann, 2006). However, in those studies parents were requested to detect truths and lies in children in general, not in their own children. Morency and Krauss (1982) found that parents were better than other adults at detecting their own child's deception.

Several studies have investigated whether it is more difficult to discriminate between truths and lies by observing children's faces and body movements than by listening to their speech or reading a transcript of their speech (Ball & O'Callaghan, 2001; Chahal & Cassidy, 1995; Feldman & White, 1980; Shennum & Bugental, 1982; Talwar & Lee, 2002). Chahal and Cassidy found no difference in lie detection performance when observers could either see the children's faces or their body. In the other studies a similar trend emerged as was found with adults. Observers were more accurate in detecting children's truths and lies by listening to their voices and speech or by reading a transcript of their speech than by observing their behaviour.

Some studies have examined whether children become better lie detectors as they get older. One might expect this to be the case. DePaulo, Jordan *et al.* (1982) suggest that with age, children obtain more cultural, social, and interpersonal knowledge. This may help them to realise that certain affects, events, or experiences described by other people are unlikely to occur in the way they are described. This life experience argument could perhaps explain why sexually abused children make better lie detectors than those who have not been abused (Bugental, Shennum, Frank, & Ekman, 2001). Also, it sounds reasonable that role-taking skill is related to ability to detect lies (Feldman, White, & Lobato, 1982; Saarni & von Salisch, 1993). Understanding the circumstances under which the senders made their statements should help the observer to interpret whether the other person's behavioural reactions are appropriate given the circumstances.

Although the research is too limited to give a definite answer, all studies addressing this issue have found that children's ability to detect truths and lies increases with age (DePaulo, Jordan *et al.*, 1982;

Feldman *et al*., 1982; Morency & Krauss, 1982; Rotenberg, Simourd, & Moore, 1989). DePaulo, Jordan *et al*. (1982), who studied this issue thoroughly, found a nearly perfect linear relationship between increasing age and improvement in ability to detect deceit. Feldman *et al*. (1982), who investigated children's ability "to put themselves into the position of another person" prior to the lie detection task, found a positive correlation between this ability and the ability to detect truths and lies.

Summary

Laypersons are not good at detecting truths and lies told by children. Being a parent does not facilitate lie detection in children, but parents may be better than other adults in detecting their own child's deceit (although this has been examined in only one study). Children's truths and lies are easier to detect when observers listen to what the child says than by observing their behaviour, and children's ability to detect deceit improves as they get older.

PROFESSIONAL LIE CATCHERS' ABILITY TO DETECT DECEPTION IN ADULTS

It could be argued that university students (who are typically the observers in lie detection studies) do not habitually detect deception. Perhaps professional lie catchers, such as police officers or customs officers, can outperform laypersons? It may be that their experience in interviewing people and catching liars has a positive influence on their skills to detect deceit. Both professionals themselves and laypersons think that professional lie catchers are better at identifying truths and lies than laypersons (Garrido, Masip, & Herrero, 2004).

Professional lie catchers have participated as observers in numerous studies. As was the case with laypersons, they saw fragments of people they did not know and the only information available to them was the nonverbal and verbal behaviour displayed by the senders in those fragments. Table 6.3 shows the studies published in English that I am aware of. Apart from the accuracy rates, it also shows the professional background of the participants. It can be seen that this group mostly comprised police officers. The top part of the table refers to studies where professionals judged adult senders, whereas the lower part of the table deals with the professionals' ability to detect deceit in children. In three adult studies the "professionals" were prisoners. I will discuss these findings first.

Table 6.3 Accuracy rates of professional lie catchers judging adults[1]

	Accuracy rate (%)		
	Truth	Lie	Total
G. D. Bond, Malloy, Thompson, Arias, & Nunn (2004, prisoners)[2]	35	69	52
G. D. Bond, Malloy, Arias, Nunn, & Thompson (2005, prisoners)	51	67	59
Hartwig, Granhag, Strömwall, & Andersson (2004, prisoners)	42	89	65
DePaulo and Pfeifer (1986, federal law enforcement)[3]	64	42	53
Ekman & O'Sullivan (1991, Secret Service)			64
Ekman & O'Sullivan (1991, federal polygraph examiners)			56
Ekman & O'Sullivan (1991, police officers)			56
Ekman, O'Sullivan, & Frank (1999, CIA)	66	80	73
Ekman, O'Sullivan, & Frank (1999, sheriffs)	56	78	67
Ekman, O'Sullivan, Frank (1999, law enforcement)	54	48	51
Garrido, Masip, & Herrero (2004, police officers)	26	69	48
Gozna & Forward (2004, police officers)			48
Hartwig, Granhag, Strömwall, & Kronkvist (2006, police detectives)			56
Hartwig, Granhag, Strömwall, & Vrij (2004, police detectives)			57
Kassin, Meissner, & Norwick, (2005, police officers)	64	32	48
Köhnken (1987) (police officers)	58	31	45
* Mann & Vrij (2006)[4]	67	70	69
* Mann, Vrij, & Bull (2004, police officers)	63	66	65
* Mann, Vrij, & Bull (2006, police officers)	73	69	71
* Mann, Vrij, Fisher, & Robinson (2007, police officers)	60	70	65
Masip, Garrido, & Herrero (2003b, police officers)[5]	42	62	52
Meissner & Kassin (2002, law enforcement)			50
Porter, Woodworth, & Birt (2000, parole officers)	20	60	40
Vrij (1993) (police detectives)	51	46	49
Vrij, Akehurst, Brown, & Mann (2006, teachers, social workers, police officers)	58	63	61
Vrij & Graham (1997, police officers)			54
* Vrij & Mann (2001a, police officers)	70	57	64
* Vrij & Mann (2001b, police officers)		51	
Vrij, Mann, Kristen, & Fisher (2007, police officers)[6]	63	38	50
* Vrij, Mann, Robbins, & Robinson (2006, police officers)[7]	70	73	72
Vrij, Mann, Fisher, Leal, Milne, & Bull (in press, police officers)	50	42	46
Total accuracy scores[8]	**56.35**	**56.11**	**55.91**

Notes: *Judging truths and lies in real-life high-stakes situations; [1]Only control conditions are included; [2]Pre-probing analyses only; [3] Experienced and inexperienced officers combined; [4]Conditions 1 and 3 combined; [5]Judgements at moment 3, results for order of presentation combined; [6]Three interview conditions combined; [7]All four tests combined; [8]Results from prisoners were not taken into account when calculating these accuracy scores

Prisoners

Both of Bond *et al.*'s prisoners' studies were carried out in a prison in the United States and Hartwig *et al.*'s study was conducted in a Swedish prison. The results of the three studies were similar. Prisoners were much better at detecting lies than at detecting truths. Lie accuracy rates ranged from 67% to 89%, whereas truth accuracy rates ranged from 35% to 51%. In all three studies laypersons (undergraduate students) also participated as observers (see Appendix 6.1), making it possible to directly compare the lie detection performance of prisoners and laypersons. Prisoners outperformed the laypersons in terms of overall accuracy in all three studies. Prisoners' better ability to distinguish between truths and lies is perhaps not surprising given that they are more knowledgeable than laypersons about how liars behave (Chapter 5).

In contrast to laypersons, who showed a truth-bias in two studies and no bias in the third study, prisoners showed a lie-bias in all three studies and judged most of the fragments as dishonest. A lie-bias in prisoners is perhaps not surprising, as deception is probably more common in the criminal world, and for criminals to be credulous towards fellow criminals may result in being deceived by them. If these contacts are business contacts it may result in being exploited, and having a suspicious mind in such circumstances is the best strategy.

Professional Lie Catchers: Total Accuracy Rates

In the remaining 24 studies (representing 28 samples) listed in Table 6.3 professional lie catchers, mostly police officers, participated. Their total accuracy rates ranged from 40% to 73%, resulting in an average total accuracy rate of 55.91%.[10] This is comparable to the total accuracy rate obtained with laypersons (Appendix 6.1). In eight studies the lie detection abilities of police officers and laypersons were compared (DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991; Garrido & Masip, 2001; Garrido *et al.*, 2004; Kassin, Meissner, & Norwick, 2005; Masip, Garrido, & Herrero, 2003b; Meissner & Kassin, 2002; Vrij & Graham, 1997). In seven of those eight studies no differences were found in total accuracy rates between police officers and laypersons, and in one study laypersons outperformed police investigators (Kassin *et al.*, 2005). In other words, not one single study has shown that police officers are superior to laypersons in discriminating between truth tellers and liars.

[10] For this analysis, the scores for the different samples in Ekman and O'Sullivan (1991) were averaged, as were the scores for the different samples in Ekman, O'Sullivan, and Frank (1999).

In two studies Ekman and his colleagues tested the lie detection skills of several groups of professional lie catchers. They concluded that some groups appear to be better than others in catching liars. In their study in 1991 members of the Secret Service outperformed laypersons, although their 64% accuracy rate is still moderate. In 1999, they identified a few more groups with superior lie detection skills, particularly members of the CIA who obtained a 73% total accuracy rate.[11] Ekman and O'Sullivan (1991) provided two reasons why secret agents make better lie detectors. First, many Secret Service agents have done protection work, guarding important government officials from potential attack. This work included scanning crowds and in those scanning tasks they have to rely on nonverbal cues. Perhaps their experience in relying on nonverbal cues made them better at spotting nonverbal cues to deceit in the deception task. Second, Secret Service officials believe that most people they deal with are telling the truth, in contrast to, for example, police officers who often believe that suspects are lying. The Secret Service therefore deals with a low base rate of lying and may be more focused on nonverbal and verbal signs of deceit. For police officers, who presume that many people are lying to them, detecting deceit may not be a fruitful option. They may rely more on searching for evidence to support or discredit the suspects' claims.[12]

Truth and Lie Accuracy Rates and Lie-Bias

It is only when ability to detect truths and lies are examined separately that differences between professional lie catchers and laypersons emerge. Laypersons were better at detecting truths than lies, but this trend is not found in professional lie catchers. Their truth accuracy rates ranged from 20% to 73% and their lie accuracy rates varied from 31% to 80%. In nine samples they were better at detecting truths than lies, and in the other 10 samples they were better at detecting lies than truths. The average truth accuracy rate for professional lie catchers was 56.35% and the average lie accuracy score was 56.11%.[13] Compared to laypersons, professionals seem to be somewhat better in detecting lies and somewhat worse in detecting truths. This is, at least

[11] It may not all be as positive as it sounds. Bond (in press) suggests that Ekman, O'Sullivan, and Frank (1999) failed to report that the officers who achieved a 73% accuracy rate on one lie detection test achieved lower scores on two other tests.

[12] More recently, O'Sullivan and Ekman (2004) and O'Sullivan (2005, 2007) claim to have discovered 29 individuals with superior lie detection skills after having tested over 12,000 professionals for their expertise in lie detection. They call them "wizards". See Bond and Uysal (2007) for a statistical critique of the evidence for this claim.

[13] For this analysis, the scores for the different groups in Ekman, O'Sullivan, and Frank (1999) were averaged.

in part, because professional lie catchers do not appear to show a truth-bias. In fact, many studies have demonstrated that police officers show a lie-bias, judging the majority of the fragments they were exposed to as being deceptive.[14] Moreover, in every study where both professional lie catchers and laypersons have participated, the professional lie catchers were more inclined to judge a fragment as deceptive than laypersons (Garrido *et al.*, 2004; Kassin *et al.*, 2005; Masip, Alonso *et al.*, 2005; Masip *et al.*, 2003b; Meissner & Kassin, 2002). Finally, the longer police officers work in the police force, the more pronounced the lie-bias becomes (Masip, Alonso *et al.*, 2005; Meissner & Kassin, 2002).[15] Perhaps socialisation within the police force increases police officers' suspicion, which subsequently leads to making more lie judgements (Masip, Alonso *et al.*, 2005).[16]

Confidence

In addition to accuracy, many lie detection studies measure observers' confidence in the veracity judgement they make. In such studies a relationship between confidence and accuracy is typically not found (see DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997, for a meta-analysis). That is, confidence does not predict accuracy. Studies where professional lie catchers' and laypersons' confidence and accuracy levels are compared show an interesting trend: professional lie catchers are more confident in their veracity judgements than laypersons, but are no more accurate (DePaulo & Pfeifer, 1986; Garrido *et al.*, 2004; Masip, Garrido, & Herrero, 2004; Meissner & Kassin, 2002). Allwood and Granhag (1999) pointed out that the tendency to be confident is not unique to police officers or lie detection, but common amongst many groups of professionals in carrying out various tasks.

[14] Garrido *et al.*, 2004; Masip, Alonso *et al.*, 2005; Masip *et al.* (2003b); Hartwig, Granhag, Strömwall, & Vrij, 2004; Kassin *et al.*, 2005; Meissner & Kassin, 2002, 2004.

[15] Most studies did not reveal a relationship between experience in the police force and *accuracy* in detecting truths and lies (DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991; Meissner & Kassin, 2002; Porter, Woodworth, & Birt, 2000).

[16] There may be a cultural element in the tendency for professionals to judge people as liars. Our British studies do not show a lie-bias (Mann & Vrij, 2006; Mann, Vrij, & Bull, 2004, 2006; Mann, Vrij, Fisher, & Robinson, 2007; Vrij, Akehurst, Brown, & Mann, 2006; Vrij & Mann, 2001b; Vrij, Mann, Fisher, Leal, Milne, & Bull, in press; Vrij, Mann, Kristen, & Fisher, 2007; Vrij, Mann, Robbins, & Robinson, 2006). Perhaps this is related to police culture. British writings emphasise an ethical approach to police interviewing that has open-mindedness of the interviewer as a core aspect (e.g., Williamson, 1993). In contrast, American manuals mainly emphasise tactics that could be used to break a suspect's resistance in order to obtain confessions (e.g., Inbau, Reid, Buckley, & Jayne, 2001). Those tactics are based on an assumption of guilt of the suspect, and, as I described earlier, assuming guilt will lead to a lie-bias.

However, high confidence in one's ability to catch liars can be harmful when the confidence is unjustified (Kalbfleisch, 1992). High confidence often results in making quick decisions on the basis of limited information (Levine & McCornack, 1992; Lord, Ross, & Lepper, 1979). Imagine the following situation. Someone is ready to go on holiday by car, but the weather is particularly bad at the time that the person wants to leave. In such a situation, an insecure driver would probably gather more information about the circumstances on the road than an experienced driver, and therefore would make a more well-considered decision about whether to drive or not. For example, the insecure driver will listen to the weather forecast to find out how the situation will develop. The experienced driver will probably rely upon inadequate heuristics such as "The weather won't stay that bad for very long" or "The weather is notoriously bad here, but it will probably improve when I get nearer to my holiday destination". In a similar vein, high confidence in lie detection skills may also lead to using heuristics when identifying truth tellers and liars. The use of heuristics in detecting deception is error prone. As mentioned in Chapter 3, the relationship between deception and nonverbal behaviour is too complicated to be approached with simple decision rules.

High confidence can be harmful for other reasons as well. It could make investigators keen to attempt detecting lies via judging someone's demeanour, which could occur at the expense of searching for physical evidence (Colwell, Miller, Lyons, & Miller, 2006). High confidence is also likely to reduce motivation to learn more about lie detection, as investigators may consider themselves already knowledgeable about the topic. An unwillingness to learn more about lie detection is undesirable given professional lie catchers' typical moderate performance in this task. In addition, as I will discuss in more detail in the Discussion of this chapter, if a police detective is confident that a suspect is lying, the detective may submit the suspect to persuasive interrogation techniques to obtain a confession. If the suspect is innocent, this may lead to a false confession. Finally, high confidence may have consequences when information is presented in court. Research has indicated that jurors are particularly influenced by how confident witnesses are (Cutler, Penrod, & Dexter, 1990; Cutler, Penrod, & Stuve, 1988; Lindsay, 1994), suggesting that jurors are more likely to believe a police officer who expresses with confidence that a suspect's demeanour revealed that the suspect was lying than an officer who does not express such confidence.[17] Being

[17] This argument assumes that professional lie catchers present veracity assessments based on observing suspects' demeanour as evidence in court. I do not know how often this happens, but I am aware of one case in the UK.

guided by high confidence is problematic if the highly confident officer is not accurate.

Stakes

As I already mentioned, in lie detection studies observers are typically exposed to senders who lied or told the truth for the sake of the experiment. One could argue that such studies do not accurately measure police officers' ability to detect deceit, for example, because the stakes for senders in such studies are much lower than the stakes for suspects during their police interviews. Given the problems of introducing high stakes in laboratory settings (Chapter 3) one could argue that the only valid way to investigate police officers' true ability to detect deception is to examine their skills when they detect truths and lies told in real-life criminal investigation settings. We did exactly this in a series of lie detection experiments. All studies in Table 6.3 marked with a '*' are studies in which police officers were asked to distinguish between truths and lies told in real-life high-stakes situations. In several studies, police officers were shown videotaped fragments of police interviews with suspects accused of crimes such as murder, arson, and rape (Mann & Vrij, 2006; Mann, Vrij, & Bull, 2004, 2006; Mann *et al.*, 2007; Vrij, Mann, Robbins, & Robinson, 2006). The fragments were derived from Mann, Vrij, and Bull's (2002) analysis of real-life police–suspect interviews described in more detail in Chapter 3. Those studies showed relatively high total accuracy rates (ranging from 65% to 72%) suggesting that police officers are reasonably accurate at detecting high-stakes truths and lies. However, one should be cautious when drawing this conclusion. In another study (Vrij & Mann, 2001b), we exposed police officers to videotaped press conferences of people who were asking the general public for help in finding their missing relatives or information about the murderers of their relatives. They all lied during these press conferences and they were all subsequently found guilty of having killed the missing person themselves. The police officers performed at the level of chance in this study (51% accuracy rate). Finally, we showed police officers fragments of the interview with a convicted murderer (Vrij & Mann, 2001a), whose case was described in detail in Chapter 3. Police officers classified a relatively high number of truths correctly (70%), but their lie detection performance was moderate (57%).[18] The findings

[18] It was already known that people had a tendency to believe this man. From conversations with police detectives who were involved in the extensive investigation, I learned that several detectives had believed that he was innocent. Also, the man had a criminal history and had been imprisoned before for statutory rape. Although he could have been kept in prison indefinitely, he was released because he was considered to no longer be a

Table 6.4 Accuracy rates of professional lie catchers judging children[1]

		Accuracy rate (%)		
	Age of sender	Truth	Lie	Total
Chahal & Cassidy (1995, social workers)	8	63	70	67
Crossman & Lewis (2006)	3–7	42	55	49
Hershkowitz, Fisher, Lamb, & Horowitz (2007)[2]		95	24	60
Jackson & Granhag (1997, barristers)	11–12	54	32	43
Leach, Talwar, Lee, Bala, & Lindsay (2004, exp 1, police officers)	3–11			44
Leach, Talwar, Lee, Bala, & Lindsay (2004, exp 1, customs officers)	3–11			49
Vrij, Akehurst, Brown, & Mann (2006, teachers, social workers, police officers)	5–6	61	63	62
Westcott, Davies, & Clifford (1991, physiotherapists, education and occupational psychologists, counsellors, career advisors)	7–11	67	53	59

Notes: [1]Only control conditions are used; Hershkowitz *et al.* is a field study, the other studies are laboratory studies; [2]Protocol condition; age of the children is not reported

overall thus suggest that in high-stakes situations police officers will still frequently make errors in truth/lie detection.

Professional Lie Catchers' Ability to Detect Deception in Children

Table 6.4 shows the studies published in English that I am aware of in which professional lie catchers' ability to detect children's truths and lies has been tested. Although only eight studies (representing nine samples) have been carried out with children[19] the findings are similar to those obtained with adults. That is, the accuracy rates ranged between 43% and 67% which is comparable to the accuracy rates range achieved in adult studies. In addition, in studies where a comparison was made between professionals and laypersons (all studies except Hershkowitz, Fisher, Lamb, & Horowitz, 2007; and Westcott, Davies, & Clifford, 1991), professionals typically achieved the same lie detection performance as laypersons. Finally, despite being no more accurate

threat to society. After his release he revived his deviant sexual activities but now also went on to kill his victims in order to destroy the evidence and minimise his chances of being caught again.

[19]The eighth study, Goodman, Batterman-Faunce, Schaaf, & Kenney (2002), did not present accuracy rates and is therefore not presented in Table 6.3.

than laypersons, professionals were more confident in their decisions than laypersons (Leach, Talwar, Lee, Bala, & Lindsay, 2004).

Summary

Professional lie catchers can detect truths and lies with somewhat higher accuracy than could be expected by flipping a coin. Perhaps with the exception of secret agents, professionals do not outperform laypersons in this task. Compared to laypersons, professional lie catchers are more inclined to judge someone as dishonest and are more confident in the veracity judgements they make.

INDIVIDUAL DIFFERENCES IN THE ABILITY TO IDENTIFY TRUTHS AND LIES

Although most lie detection studies have revealed accuracy rates of between 50% and 60%, there are individual differences between observers within studies. For example, in our lie detection experiment (Mann *et al.*, 2004) where police officers assessed videotaped police–suspect interviews with murderers, rapists, and arsonists, large individual differences were found, with accuracy rates for individual officers varying from a low 30% to a very high 90% (achieved by three officers, Mann, 2001). How can such individual differences in lie detection studies be explained? First, obtaining a high or low accuracy score may be, in part, a matter of bad or good luck. In our experiment, it is possible that the officer who obtained a 30% accuracy rate will have a higher accuracy rate in a second study, and it is also possible that the three officers who achieved 90% accuracy rates will obtain lower accuracy rates in a follow-up study. Indeed, in a study where we exposed our lie detectors to four lie detection tests rather than to a single lie detection test, none of the police officers obtained consistently very low or very high accuracy rates (Vrij, Mann, Robbins, & Robinson, 2006). Despite this, after four tests some officers were still better than others (accuracy rates ranged from 62% to 82%), indicating that stable individual differences might exist.[20] Also, others have found that the ability to detect lies in one test is, to some extent, positively related to being able to detect lies in another test (Edelstein *et al.*, 2006), again suggesting that stable individual differences exist.

[20] Alternatively, it could be that if even more tests would be carried out, differences between individuals will further decline and eventually cease to exist. This is the view of Bond and DePaulo (2007).

Research has been carried out which attempts to unravel why some people are better lie detectors than others. Two factors that do not explain individual differences have already been discussed. Lie detection ability is unrelated to profession of the lie detector (perhaps with the exception of secret service agents), neither is it related to confidence in one's ability to detect lies. Other characteristics that are unrelated to lie detection are years of service as a professional lie detector (see Aamodt & Custer, 2006 for a review); and age of the lie detector (Aamodt & Custer, 2006; Vrij & Mann, 2005). This section reviews other possible relationships. I begin with discussing the relationship between gender of the lie detector and ability to identify truths and lies, followed by the relationship between lie detection skill and the personality of the lie detector, the cues lie detectors use when they detect deceit, lie detector's familiarity with the communication style of the sender, and lie detector's motivation.

Gender Differences in Lie Detection Skills

Women are superior to men in interpreting other people's nonverbal behaviour. That is, women are better than men in understanding the messages that others purposefully convey to them (Hall, 1979, 1984; Rosenthal & DePaulo, 1979). Women take more time in observing and interpreting nonverbal cues than men and use more cues in their decision-making process (Hurd & Noller, 1988). They have the biggest advantage over men in reading facial expressions, and are more accurate in interpreting the facial expressions of people who have nothing to hide (DePaulo, Epstein, & Wyer, 1993). Hall (1979) offered an accommodation explanation for these findings. She compared women's superiority over men in understanding nonverbal messages in 11 countries. She found that women were most superior in those countries where they seemed to be most oppressed, for example in countries where a small proportion of women were in higher education. Women's sex role in many societies implies that they, more than men, have to accommodate themselves to others. Being able to interpret someone's nonverbal behaviour could be a valuable skill to achieve this.

Although women are superior to men in reading nonverbal messages, they are no better than men in detecting truths and lies in strangers (DePaulo, Wetzel *et al.*, 2003; DePaulo, Epstein, & Wyer, 1993; Hurd & Noller, 1988; Manstead, Wagner, & MacDonald, 1986; Porter, Woodworth, McCabe, & Peace, 2007). Women are, however, less suspicious than men and are more inclined to believe that they are being told the truth (DePaulo, Epstein, & Wyer, 1993). In other words, women seem to lose their advantage over men in reading nonverbal behaviour in lie detection tasks and show a truth-bias when assessing strangers. The following explanation sounds reasonable. Women are better than men in decoding the information someone *wants to convey*. During deception, however, liars try to *hide* their true feelings and thoughts. When detecting lies, observers should not examine what people *want to convey* but should look at what they *try to conceal*. Perhaps, when women try to detect lies in strangers, they are too affected by what that person tries to convey, which will result in errors in lie detection and in a truth-bias.

Women, however, do no appear to lose their advantage over men in reading nonverbal behaviour when they attempt to detect truths and lies in people they know, such as their romantic partners or friends. McCornack and Parks (1990) found that women were superior to men in detecting lies in their romantic partners. In another study, women, but not men, could tell whether their romantic partner just received disappointing news, whereas neither men nor women could tell whether a stranger just received disappointing news (DePaulo, Wetzel *et al.*, 2003). Finally, female friends were more accurate at detecting each other's lies after being friends for six months than during the first month of their friendship, whereas male friends became no more insightful about each other's lies over the six-month period (Anderson *et al.*, 1999).

In summary, the findings suggest that women are better lie detectors than men, but only if they know the person whose lies they try to detect. Perhaps women are more aware then men of the natural truthful behaviour and speech of their friends and partners and so can turn this to their advantage when attempting to detect lies by comparing this truthful baseline demeanour with the demeanour under investigation. Alternatively, perhaps men and women are equally aware of the truthful demeanour of their friends and partners, but women notice subtle differences between this truthful demeanour and their friends' and partners' lies that men fail to notice.

Personality of the Observer and Lie Detection Skills

Several researchers have examined whether the observer's personality influences his or her ability to detect truths and lies. DePaulo and Tang (1994) were perhaps the first to investigate this. They examined the personality trait *social anxiety*. People high in social anxiety suffer from low self-esteem, are nervous during social interactions, and typically believe that they do not come across very well. DePaulo and Tang argued that for several reasons observers high in social anxiety would be less successful at detecting truths and lies than observers low in social anxiety. Nervousness could lead to narrowing of the focus of attention to a relatively small number of cues. This means that socially anxious observers could become too focused on a limited set of cues and fail to notice some important cues to deceit. Also, socially anxious people are likely to be distracted by thoughts unrelated to the task such as feelings of ineptitude and concerns about losing status. Those distractions could again lead to missing some important cues to deceit.

Moreover, the combination of attempting to detect lies and task-irrelevant worries put high demands on the observer's working memory. When the task becomes too demanding it may impair capability to process information that is relevant to the lie detection task. Finally, even when socially anxious observers do process information adequately and do spot some important cues, they may misinterpret them. Socially anxious people often have a defensive approach to social interactions. Their worries about the possibility of creating a bad impression on others may result in interacting safely to try to prevent this from happening. One safe approach is to read the cues in the way the sender would prefer them to be read. DePaulo and Tang's study showed that observers high in social anxiety were indeed worse at discriminating between truths and lies than observers low in social anxiety, but in other studies no differences were found between observers who were high or low in social anxiety (Vrij & Baxter, 1999; Vrij, Harden, Terry, Edward, & Bull, 2001).

Other researchers looked at the relationship between *self-awareness* and the ability to detect deceit. Self-awareness could be defined as a state of self-directed attention. People high in self-awareness, sometimes labelled *private self-consciousness*, tend to agree with statements such as "I reflect about myself a lot", "I am generally attentive to my inner feelings" and "I am constantly examining my motives". Since self-awareness gives insight into one's own mind, it could be that it also provides insight into what is going on in someone else's mind. This ability to mind read could facilitate lie detection. Indeed, positive relationships have been found between self-awareness/private self-consciousness and the ability to distinguish between truths and lies (Johnson, Barnacz, Constantino, Triano, Shackelford, & Keenan, 2004; Malcolm & Keenan, 2003).

Private self-consciousness is related to *introversion*. The introvert is generally oriented toward the internal world of ideas and concepts. Unlike in private self-consciousness, where the thoughts and reflections solely deal with the self, the internal world of introverts also covers issues other than the self. Introverts' reflections of themselves may be beneficial for lie detection in the same way as it was beneficial for those who score high in self-awareness. We did not find a relationship between introversion and the ability to distinguish between truths and lies (Vrij & Baxter, 1999; Vrij, Harden *et al.*, 2001), but O'Sullivan (2005) reported that most of the people who she claims have exceptionally good lie detection skills (the wizards, see footnote 12) seem to be introverts.

People high in self-awareness and introverts may be good at reading others' minds, good *actors* may be good at reading someone else's behaviour. I already discussed in previous chapters that liars sometimes start acting to suppress cues of nervousness and cognitive load that they do not wish to exhibit and to display the behaviour they believe appears credible. Perhaps good actors are superior at noticing whether someone is showing "natural" behaviour (e.g., truth telling) or that someone is "acting" (e.g., lying). In one study we found that good actors were better at distinguishing between truths and lies than poor actors but we could not replicate this finding in a second study (Vrij, Harden *et al*., 2001). In summary, although the findings are mixed and not conclusive, there is ample evidence that people who are high in self-awareness, introverted and good actors make relatively good lie detectors whilst socially anxious people make relatively poor lie detectors.

A different issue is whether the personality of the lie detector is related to confidence in lie detection skills. People high in social anxiety, introverts and shy people are less confident in social interactions than people low in social anxiety, extraverts and people who are not shy. It sounds reasonable that lack of confidence in social interactions will make people also less confident in judging social interactions, and this could mean that people high in social anxiety, introverts and shy people are less confident in their ability to detect deceit than their counterparts. Research supports this assumption (Vrij & Baxter, 1999; Vrij, Harden *et al*., 2001). I already explained that not being too confident is probably beneficial in lie detection given the difficulty of the lie detection task.

Relationship between Cues Used by the Observer and Lie Detection Skills

It sounds plausible that being knowledgeable about diagnostic cues to deception makes someone better at detecting truths and lies. James Forrest and his colleagues found exactly this (Forrest, Feldman, & Tyler, 2004). They first requested participants to fill out an 18-item "beliefs about cues to deception" questionnaire, and then asked them to take part in a lie detection experiment. The more accurate the participants were in answering the beliefs about cues to deception questions, the better they performed at the lie detection task.

In most other studies the relationship between cues used by observers and lie detection ability is examined in a different way. In most studies, observers are either asked before the lie detection task which cues they pay attention to when they attempt to detect deceit, or they are asked after each veracity judgement on which cues they based their decision. In our study where we showed police officers videotaped fragments of police interviews with murderers, rapists and arsonists, several relationships occurred between cues reported by the officers and their accuracy in truth and lie detection (Mann *et al.*, 2004). First, good lie detectors mentioned verbal cues (vague reply, contradictions in story, etc.) more often than poor lie detectors. Second, the more visual cues (gaze aversion, posture, movements, etc.) participants mentioned, the lower their accuracy became. Particularly, police officers who mentioned that liars look away and fidget achieved the poorest scores. In other words, those who listened carefully to what the suspects had to say were better lie detectors than those who concentrated on the suspects' nonverbal behaviour.

Other studies also revealed that listening to senders makes people better lie detectors and watching senders makes people worse lie detectors. Anderson *et al.* (1999) and Feeley and Young (2000) found that the more vocal cues (speech errors, speech fillers, pauses, voice) participants mentioned, the higher accuracy they obtained. In a study where participants attempted to detect truths and lies told by the convicted murderer described in Chapter 3, we found that participants who mentioned gaze aversion and fidgeting as cues to deceit achieved the lowest accuracy scores (Vrij & Mann, 2001a). Also Porter, Woodworth, McCabe, & Peace (2007) found that the more visual cues the participants reported, the worse their ability to distinguish between truths and lies became.[21]

Ekman and his colleagues, however, demonstrated that paying attention to visual cues could be useful in detecting deceit. Frank and Ekman (1997) reported that good lie detectors were better at spotting brief facial expressions of emotion than poor lie detectors. Ekman and O'Sullivan (1991) found that participants who mentioned both vocal/verbal and visual cues obtained higher accuracy rates than participants who just mentioned vocal/verbal or visual cues.

In summary, virtually all studies showed that in order to detect lies, listening carefully to what is said is necessary, and that merely paying attention to behaviour hampers lie detection. Those findings are in alignment with findings reported above that observers who can only see a sender perform worse than observers who can only hear, or who can both see and hear a sender. In contrast, police manuals often focus on visual cues to deception (Chapter 5). For example, Inbau and

[21] Porter, Woodworth, and Birt (2000), however, found that reporting body cues led to improved accuracy.

his colleagues (2001) suggest that liars display a variety of visual cues, including gaze aversion, unnatural posture changes, self-self-adaptors, and placing the hand over the mouth or eyes when they speak. We measured the effectiveness of using these "Inbau-cues" in our lie detection study (Mann, Vrij, & Bull, 2004). We counted the number of Inbau cues that the police officers mentioned and related this to their performance on the lie detection task. We found that the more Inbau-cues the police officers mentioned, the worse they became at distinguishing between truths and lies. Moreover, in their experiment Kassin and Fong (1999) taught some observers the visual cues that Inbau *et al.* discuss in their manual. Their performance on a subsequent lie detection test was worse than the performance of untrained participants. In other words, endorsing the information about visual cues to deception discussed in Inbau *et al.*'s (2001) manual is counterproductive and makes people worse lie detectors. Other police manuals report visual cues similar to those reported by Inbau *et al.* (Chapter 5), and it is likely that the same applies to those manuals. Endorsing the information about visual cues reported in those manuals is likely to make people worse at lie detection.

Familiarity with the Communication Style of the Liar

People become better lie detectors when they are familiar with the communication style of the liar. For example, people have a more open communication style when they speak with attractive people than when they talk to unattractive people. This means that attractive and unattractive people are used to different communication styles (DePaulo, 1994). DePaulo, Tang, and Stone (1987) examined whether this has an impact on the ability to detect truths and lies. Senders were asked to tell truths and lies to attractive and unattractive conversation partners. These statements were videotaped and presented to attractive and unattractive observers. On the videotape, only the senders were visible, not the attractive and unattractive people to whom the lies were told. The findings revealed that attractive observers were better at detecting truths and lies that were told to attractive people, whereas unattractive observers were more accurate in detecting truths and lies when they were told to unattractive people. Hence, attractive and unattractive people are better in detecting lies when they are spoken to in a communication style that they are familiar with.

A similar pattern emerges when lie detectors try to detect truths and lies that are told by fellow native residents or by people residing in a different country. Observers are more able to detect truths and lies told by senders from their own country than by foreign senders. For example, in one study observers watched video fragments that were presented without sound. American observers could detect truths and lies above the level of chance when they were told by fellow Americans senders, but failed to do so with Jordanian senders. Jordanian observers showed the reverse pattern. They were able to detect truths and lies above the level of chance in Jordanian senders but not in American senders (Bond, Omar, Mahmoud, & Bonser, 1990).

In a follow-up study, Bond and Atoum (2000) examined the ability to discriminate between truths and lies when observers could also hear what has been said. They also included Indian participants in their sample (see also Bond & Rao, 2004). This time, Americans, Jordanians, and Indians could detect above the level of chance the truths and lies told by foreigners, but only if they could both hear and see them. Just observing these foreigners' behaviour or just listening to their speech resulted in chance level performance. In other words, both vision and sound were needed to detect these foreigners' lies. Since observers could not understand what the foreigners were talking about – they did not speak these foreign languages – the sound condition presented them with vocal cues but not with verbal cues. Hearing these vocal cues was therefore a necessary requirement for lie detection.

The Motivated Lie Detector

People are not always equally motivated to detect lies. A mother may not care so much about finding out if her son is telling the truth when he tells her that he enjoyed his meal, but she may be more interested to learn the truth when he denies having taken money out of her purse. Being motivated could affect truth and lie detection in different ways. Intuitively, one would perhaps expect that motivation improves performance and that highly motivated lie detectors are better at discriminating between truths and lies than lie detectors who are not so motivated. Motivation will probably make lie detectors more attentive and therefore more likely to spot cues to deceit that they may otherwise fail to spot. However, one could also predict that increased motivation could impair lie detection ability. Psychological research has demonstrated that high levels of motivation generally facilitates performance when a task is relatively easy but impairs performance when a task is relatively difficult (Kim & Baron, 1988; Pelham & Neter, 1995; Zajonc, 1980).

To explain this phenomenon Zajonc (1980) proposed that motivation creates arousal and that arousal enhances the individual's tendency to perform their "dominant response". When a task is easy or well learned the dominant response is often correct, but when the task is complex the dominant response is often incorrect. Since lie detection is a difficult task that requires complex information processing and interpretation, motivation should thus impair rather than improve lie detection performance if Zajonc's assumption is correct. Research has supported Zajonc's assumption and demonstrated that motivation impairs ability to be able to detect truths and lies (Forrest & Feldman, 2000; Porter *et al.*, 2007). Perhaps motivation makes observers rely on the cues they believe are diagnostic indicators of deceit (e.g., dominant response), and since people's favourite cues are often non-diagnostic (Chapter 5), an impaired lie detection performance is the likely result. If this reasoning is true, increased motivation should improve the performance of good lie detectors, because motivation may help them to focus on diagnostic cues to deceit. Nobody has examined this to date.

O'Sullivan argues that there is a positive link between motivation and being a good lie detector (O'Sullivan, 2005; O'Sullivan & Ekman, 2004). However, she refers to motivation in a different way than it has been discussed so far. In the studies mentioned above, observers became motivated to perform well in a particular lie detection task, but in real life, that is, outside the context of the experiment, these observers may not always have been motivated lie detectors. O'Sullivan refers to people who are motivated to catch liars in real life. She argues that good lie detectors want to improve their ability to understand others and want to seek feedback about their performance. Regarding the latter, she describes a few wizards of lie detection who were distressed when they made an incorrect judgement in a lie detection task. They mentioned it several times, still referred to it days later and attempt to understand what they did wrong (O'Sullivan, 2005). In other words, O'Sullivan refers to people who are motivated to become more knowledgeable about deceit. It sounds reasonable that this will improve their lie detection skills because, as I discussed above, being knowledgeable about deception is likely to facilitate lie detection.

Summary

Several factors influence the ability to detect truths and lies. Females are better at this task than males when they attempt to detect deceit in their friends and romantic partners but they do not outperform men when they judge strangers. Moreover, people high in self-awareness, introverts and good actors are relatively good lie detectors, but socially anxious persons are relatively poor at detecting lies. Good lie detectors are more knowledgeable about diagnostic cues to deceit than poor lie detectors. Good lie detectors further listen carefully to what senders say or pay attention to both senders' speech and their behaviour, whereas poor lie detectors pay attention only to senders' behaviour. Lie detection is somewhat easier when lie detectors are familiar with the communication style used by the senders than when they are not. Being motivated to discriminate between truths and lies impairs performance in lie detectors who are not knowledgeable about cues to deceit. However, being motivated probably improves lie detection skills if that motivation results in becoming more knowledgeable about cues to deceit.

FACTORS AFFECTING VERACITY JUDGEMENTS

In the previous section I discussed factors that affect lie detectors' accuracy in truth/lie detection. In this section I discuss factors that affect observers' veracity judgments rather than their accuracy. By veracity judgements I mean whether an observer is inclined to judge a statement as truthful or deceptive. I already discussed one factor that influences veracity judgements: being suspicious results in making more lie judgments. In this section I discuss four more factors: personality of the sender; ethnic background of the sender; whether the observer is an active interviewer or passive; and asking questions ("probing").

Personality of the Sender

In Chapter 5 I discussed that people hold stereotypical views about how liars behave. A consequence of those views is that some people, regardless of whether they are lying or telling the truth, typically make a suspicious impression on observers whereas others typically make an honest impression on observers. Those people whose natural honest behaviour fits the stereotype of how liars behave give the impression that they are lying (dishonest demeanour bias), whereas those people whose natural behaviour fits the stereotype of how honest people behave give the impression that they are telling the truth (honest demeanour bias).[22]

The dishonest and honest demeanour biases are related to personality traits. People with a strong sense of *public self-consciousness* tend to make a less credible impression on others, regardless of whether they are telling the truth or lying. These are individuals who are overly concerned about being scrutinised by others, which makes them anxious. This anxiety subsequently shines through in their behaviour which makes them appear dishonest. *Introverts* and *socially anxious* people

[22] Feldman, Tomasian, & Coats, 1999; Riggio, 2006; Riggio & Friedman, 1983; Riggio, Tucker, & Throckmorton, 1988; Riggio, Tucker, & Widaman, 1987; Vrij, 1993; Vrij & Van Wijngaarden, 1994; Vrij & Winkel, 1992b; Zuckerman, DeFrank, Hall, Larrance, & Rosenthal, 1979.

also make a dishonest impression on others. The social clumsiness of introverts and the impression of tension, nervousness or fear that is naturally given off by socially anxious individuals (Schlenker & Leary, 1982) is interpreted by observers as indicators of deception.

In contrast, *expressive people* exude credibility, regardless of the truth of their assertions. Expressiveness or "spontaneous sending" (DePaulo & Friedman, 1998) can be defined as "the ease with which people's feelings can be read from their nonverbal expressive behaviours when they are not deliberately trying to communicate their feelings to others" (DePaulo & Friedman, 1998, p. 13). Expressiveness is related to charisma (Friedman, Riggio, & Casella, 1988). The first impression expressive people make is typically positive: they are generally well liked and often regarded as attractive (DePaulo & Friedman, 1998; Gallaher, 1992). Their spontaneity tends to disarm suspicion, which makes it easier for them to get away with their lies (Riggio, 1986). Finally, individuals who are *socially tactful and competent* make an honest impression. Such individuals are comfortable and relaxed in social interactions and are well practised in presenting themselves effectively.

Interestingly, those demeanours are not an accurate reflection of people's dispositions to lie (see also Chapter 2). For example, introverted people do not lie frequently (Kashy & DePaulo, 1996), and commit fewer crimes than extraverts (Eysenck, 1984). Furthermore, socially anxious people are less likely to persist in lying as soon as they are challenged (Vrij & Holland, 1998).[23,24]

Box 6.1 A demeanour bias test

There may be an easy way to find out whether you make a credible impression on others. People who are good at expressing basic emotions via facial expressions (happiness, anger, fear, surprise, sadness, and disgust) appear more credible than those who are not so good at this (Riggio & Friedman, 1983). Hence, try to express basic emotional facial expressions and ask others to determine which emotions you try to convey. The percentage of correct answers they give may be an indication of how credible the impression is that you project.

[23] People high in Machiavellianism often lie (Chapter 2), and do not experience much guilt when they lie (Chapter 3). One would think that this makes them good liars. Intriguingly, this is not the case. It is not easier or more difficult to detect truth and deceit in people high in Machiavellianism than in individuals low in Machiavellianism (Frank & Ekman, 2004; Manstead, Wagner, & MacDonald, 1986).

[24] See Chapter 5 for factors other than behaviour that create a demeanour bias, such as having an attractive or baby-faced appearance, and the sender's clothing.

Ethnic Origin of the Sender

Nonverbal behaviour is culturally mediated (Chapter 3). Take for example gaze aversion. Afro-American people display more gaze aversion than white American people (LaFrance & Mayo, 1976) and people from Turkey and Morocco who are living in the Netherlands show more gaze aversion than native Dutch people (van Rossum, 1998; Vrij, Dragt, & Koppelaar, 1992). It thus appears that looking into the eyes of the conversation partner is typical Caucasian behaviour that is often not displayed by non-Caucasian individuals. Differences in culture contribute to this effect. Looking into the eyes of a conversation partner is regarded as polite in Western cultures but is considered to be rude in several other cultures (Vrij & Winkel, 1991; Vrij, Winkel, & Koppelaar, 1991; Winkel & Vrij, 1990).

More culturally determined differences in nonverbal behaviour have been found. In the Netherlands, we examined the nonverbal behavioural patterns of native Dutch Caucasian and black Surinamese residents (citizens originated from Surinam, a former Dutch colony, but now living in the Netherlands) when they lied or told the truth during simulated police interviews (Vrij & Winkel, 1991). Either a Dutch or a Surinamese interviewer conducted the interviews, but this had no impact on the findings. Moreover, we did not find different cues to deception in the two ethnic groups.[25] However, we found large differences in behaviour between Dutch Caucasian and Surinamese participants, regardless of whether they were telling the truth or lying. Surinamese people made more speech disturbances, exhibited more gaze aversion, smiled more often, and made more self-adaptors and illustrators, regardless of whether they were lying or not.

This means that observers need to be careful in cross-cultural interactions and should interpret the nonverbal behaviours displayed by senders of a different ethnic origin with knowledge of the culture (Ruby & Brigham, 1997; Vrij, 1991). We obtained evidence that this does not always happen. We prepared videotapes of simulated police interviews in which native Dutch Caucasian and Surinamese senders (professional actors) participated. Different versions were made of each interview. The senders demonstrated typical "Dutch Caucasian" behaviour in one version of the interviews (e.g., showed a limited amount of gaze aversion) and typical "Surinamese" nonverbal behaviour in another version of the interviews (e.g., showed more gaze aversion). In other videotapes we manipulated the number of self-adaptors and speech disturbances the senders made. Caucasian Dutch police officers saw one version of each interview and were asked to indicate to

[25] Neither did Sitton and Griffin (1981) when they compared the nonverbal cues to deceit displayed by black and white senders in an American study (Chapter 3).

what extent the actor made a suspicious impression. The police officers found Surinamese senders and Dutch Caucasian senders equally suspicious. However, the nonverbal behaviour displayed by the senders did affect the police officers' impressions. The senders consistently made a more suspicious impression when they displayed typical Surinamese behaviour than when they exhibited typical Dutch Caucasian behaviour (Vrij & Winkel, 1992b, 1994). For example, the experiment with self-adaptors revealed that 72% of the police officers found the sender suspicious when he showed typical Surinamese nonverbal behaviour, whereas only 41% found him suspicious when he showed typical Dutch Caucasian nonverbal behaviour.

These findings could easily be explained when people's beliefs about how liars behave are taken into account. As Chapter 5 revealed, gaze aversion, self-adaptors and speech disturbances, behavioural patterns that are typical for Surinamese citizens, are all perceived as indicators of deceit. These findings therefore reveal that in interactions between non-Caucasian senders and Caucasian observers *cross-cultural nonverbal communication errors* occur, whereby nonverbal behavioural patterns that are typical for an ethnic group are easily interpreted by Caucasian observers as signs of deception.

Active Interviewer versus Passive Observer

In the lie detection studies presented so far, observers did not actually interview the senders but just observed them on videotape. Would accuracy improve if observers were to interview the senders? Police officers, prosecutors, and judges believe it would (Strömwall & Granhag, 2003b). Perhaps they think that interviewing gives observers the opportunity to ask questions that outplay the liar. It therefore probably comes as a surprise to professionals that interviewing in itself does not benefit lie detection. Several experiments have shown that the accuracy rates of actual interviewers (laypersons and professional lie catchers) are never higher, but are sometimes lower, than the accuracy rates of passive observers (Buller, Strzyzewski, & Hunsaker, 1991; Burgoon, Buller, & Floyd, 2001; Durban, Ramirez, & Burgoon, 2003; Feeley & DeTurck, 1997; Granhag & Strömwall, 2001b, c; Hartwig, Granhag, Strömwall, & Vrij, 2004; Kalbfleisch, 1994; Stiff, Kim, & Ramesh, 1992).[26]

A difference emerges between interviewers and observers in the type of veracity judgements they make. Interviewers (both laypersons and

[26] A different picture emerges when specialised interview protocols are taken into account. In Chapter 15 I will explain that certain interview techniques can enhance ability to detect deceit.

professional lie catchers) tend to believe senders more than observers do (Bond & DePaulo, 2006; Hartwig, Granhag, Strömwall, & Vrij, 2004). Perhaps senders respond to the interviewers' reactions during the interview and succeed in crafting a verbal message and nonverbal presentation style that appears credible to the interviewers. Since senders do not interact with the observers who watch their videotaped interviews, they cannot influence these observers' responses. Although this would result in low accuracy rates and a truth-bias, the explanation is not entirely satisfactory, because liars do not always succeed in making a convincing impression by attempting to do so (Chapter 3). Alternatively, perhaps senders do not try to make a convincing impression, but simply try to be liked by the interviewer. People generally have the desire to be liked and one way of achieving this is to be kind to others (Baumeister, 1982; Monahan, 1995). Perhaps interviewers tend to evaluate the senders more positively than observers do, and, in turn, interviewers may then perceive those they like as being honest.

Probing

A sender's statement may result in the interviewer reacting by asking for further explanation. Questions (probes) could be neutrally phrased ("I don't understand this, could you please explain this to me?"); positively phrased ("I do believe you, but I don't understand this. How is it possible that. . .?"); or negatively phrased ("I don't believe you, are you trying to fool me?"). Intuitively, one might think that further questioning makes truth and lie detection easier. The liar is forced to continue to speak and give more information. Obviously, the more liars speak and the more information they give, the greater the possibility that they will make mistakes and give their lies away, either via verbal cues (by contradicting themselves or by saying something which the observer knows is incorrect) or via nonverbal cues. However, several studies have shown that probing does not increase accuracy, but tends to lead to judging the other as being truthful (Bond, Malloy, Thompson, Arias, & Nunn, 2004; Buller, Comstock, Aune, & Strzyzewski, 1989; Buller, Strzyzewski, & Comstock, 1991; Levine & McCornack, 2001; Stiff & Miller, 1986). This is called the *probing-heuristic* (Levine, Park, & McCornack, 1999). The type of probing (negative, neutral, or positive) is irrelevant; all types of probing yield the same effect and benefits liars.

Levine and McCornack (2001), however, found that the probing effect has a limitation. Probing did result in a truth-bias when observers were merely asked to judge the veracity of each sender, but had no effect when, prior to the lie detection task, observers were informed about some diagnostic and non-diagnostic cues to deception and were instructed to solely focus on the diagnostic cues while making their veracity judgements. Levine and McCornack argued that receiving no instructions resulted in heuristic processing whereas providing instructions resulted in active information processing. Apparently, the probing effect does not occur when observers are involved in active information processing. This may explain why Granhag and Strömwall (2001b, c) did not find a probing effect in their experiments where they showed observers three interviews with the same suspect. In these studies active processing took place because observers were trying to detect contradictions between statements.

Summary

Several factors influence the veracity judgements observers make. Some individuals (people high in public self-consciousness, introverts, socially anxious people) make a dishonest impression on observers regardless of whether they telling the truth or lying, due to the behaviour they naturally show. Also non-Caucasian senders make a dishonest impression on Caucasian observers for this same reason. In contrast, people who are expressive and socially tactful exert an honest impression on observers because of the behaviour they naturally display. Being an active interviewer rather than an observer results in a tendency to believe senders, as does probing the sender.

DISCUSSION

Laypersons' Ability to Detect Deceit

This chapter examined how good laypersons are in detecting truths and lies told by strangers, friends, romantic partners, and children if they are not given any background information about the senders or their statements so that the only source of information available to them is the nonverbal and verbal behaviour displayed by these senders. In such a situation laypersons perform just above the level of chance, and are not really any better at distinguishing between truths and lies told by their friends/partners or children than they are at distinguishing between truths and lies told by strangers. People may react in disbelief and argue that they have caught their partners or children out on numerous occasions. However, they probably did so by comparing a statement with evidence they knew (the senders contradicted facts or statements of others, etc.) rather than by simply judging the demeanour of partners or children (Park *et al.*, 2002). Also, what does detecting numerous lies actually say? People do not know how many times their partners and children have lied to them. It could be that the numerous lies they did detect still represents a small proportion of the total number of lies told to them.

I further explained that laypersons have a truth-bias and tend to judge the majority of statements they encounter as truthful. The truth-bias is stronger in friends and relatives than in strangers; stronger when interacting with senders rather than observing them; and stronger when observers question (probe) senders than when they do not question them. This makes the truth-bias perhaps more profound in daily life than research suggests because in daily life people often talk to friends or relatives and interact with them.

The intriguing consequence of the truth-bias is that people will judge most of their social interactions correctly in daily life, despite being poor lie detectors (Levine, Kim, & Park, & Hughes, 2006; Park & Levine, 2001). The explanation for this is that most statements in daily life are truthful (Chapter 2). Let's give an example in percentages. Suppose that a woman held 10 conversations with her husband during a day and that he lied to her twice. And let's further suppose that she believed all of his stories. In this case, the woman has correctly classified most of her husband's stories 80% (8 out of 10). However, in terms of lie/truth discrimination her performance was poor. She correctly classified 100% of his truths (8 out of 8) and 0% of his lies (0 out of 2), giving her a total accuracy score of 50% (i.e., truth and lie accuracy combined), which is what could be expected by chance alone. Whether judging this high percentage of social interactions correctly (80%) is satisfactory depends on the nature of the lies that remain undetected. I explained in Chapter 2 that the majority of lies are harmless or even suit the observers and their relationships with others. From this perspective, a credulous strategy whereby being lied to, and the other side of the coin – deceiving others – is accepted could be advantageous. However, some lies are harmful to observers or their environment (relatives, work, etc.). In that case they are better off by changing tactics and become less credulous. The remaining part of this book gives information about what to do in order to detect lies.

Professional Lie Catchers' Ability to Detect Lies

In lie detection studies various groups of professional lie catchers have participated, albeit mostly police officers. Those studies showed that most groups of professionals perform only just above the level of chance when they attempt to detect truths and lies in people they do not know on the basis of nonverbal and verbal behaviour. Their ability to distinguish truths from lies is comparable to that of laypersons. Three reasons come to mind as to why many professionals are not better than laypersons.

First, perhaps many professionals are not practised enough in detecting truths and lies via observing someone's nonverbal and verbal behaviour, and perhaps, like laypersons, they mostly detect truths and lies by comparing statements with available evidence. This may explain why secret agents, who are experienced in scanning crowds, seem to be better lie detectors than laypersons and other groups of professionals when observing someone's demeanour.

Second, perhaps professionals lack proper feedback in the decisions they make (DePaulo & Pfeifer, 1986). As already discussed in Chapter 5, for feedback to be adequate, it needs to be frequent, reliable, and immediate. In other words, observers should be informed immediately after every interaction with a person whether that person was lying or not. Adequate feedback is difficult to achieve for most groups of professionals. Take for example police officers who interview suspects and witnesses. Whether these suspects and witnesses told the truth or lied may never come out, or may only be revealed long after the interview took place. In that respect, other professionals, such as customs officers, could receive more adequate feedback if they wished, by stopping people randomly and deciding, before searching, whether each of these persons is trying to smuggle. Their subsequent search should then confirm whether or not their intuitions were correct.

Third, perhaps professional lie catchers lack training or receive inadequate training in lie detection. I do not know what kind of training professional lie catchers receive in lie detection around the world, and indeed they may not receive any training at all. However, reading police manuals about this topic makes me pessimistic about the quality of such training programmes where they do exist.

Apart from having difficulty in detecting truths and lies two more findings emerged. First, professionals are inclined to express confidence in their ability to detect truths and lies and are typically more confident than laypersons. Being confident but not particularly skilled in a task is worrisome, amongst other reasons because high confidence often results in making quick decisions on the basis of limited information. Lie detection is a complex task, and therefore quick and shallow decision-making processes do not bode well. Second, at least in some countries professional lie catchers typically exert a lie-bias and tend to disbelieve senders.

Some factors (probing, active interviewing, and being suspicious) affect a lie-bias. Since probing and active interviewing weakens the lie-bias whereas being suspicious strengthens it, it is difficult to say whether the lie-bias sometimes obtained with professional lie catchers in laboratory research reflects real life. However, I suspect it does because I expect these factors to balance each other out. Asking questions (probing) typically makes the sender appear more credible and reduces a lie-bias only when observers process information globally (heuristically). Observers who actively search for cues to deceit (e.g., active information-processing) are not affected by probing. I assume that professional lie catchers are active rather than heuristic information processors and probing will therefore probably not affect them too much. Many professional lie catchers do actively interview people in real life and this may make them less prone to a lie-bias than laboratory research suggests. However, professionals are in real life often suspicious of the people they interview, and probably more suspicious than they are towards the senders they assess in lie detection experiments. This may make them more prone to a lie-bias in real life than laboratory research suggests. Perhaps the effects of active interviewing and being suspicious counterbalance each other, and, as a result the lie-bias obtained in laboratory research reflects real-life veracity judgements.

The consequences for innocent suspects

The picture that arises from this chapter is that when professional lie catchers attempt to pinpoint truth tellers and liars from their nonverbal and verbal behaviour, their decisions could be incorrect, lie-biased, yet made with confidence. This means that from time to time the police will interrogate innocent suspects who they mistakenly believe to be guilty. Once interrogators mistakenly believe that the innocent suspect is guilty, their incorrect belief could be strengthened through the subtle and subconscious process of mirroring.[27] The innocent suspect is likely to deny any involvement in the crime which may irritate the interrogator. Interrogators may show their irritation by displaying increased movement. Research has demonstrated that suspects respond to an increase in interviewer's movements by making more movements themselves and mirroring their behaviour. Interviewers subsequently interpret the increase in movements displayed by the innocent suspects as signs of deceit (Akehurst & Vrij, 1999).

When interrogators believe that a suspect who denies involvement in a crime is guilty, they are advised to submit the suspects to interrogation techniques such as "The Reid Nine Steps of Interrogation" (Inbau *et al.*, 2001). This is a persuasive, confrontational interrogation technique which is potentially powerful enough to make suspects

[27] See also *the chameleon effect* in Chapter 3.

confess, including those who are innocent (Gudjonsson, 2003; Kassin, 1997, 2004; 2005; Kassin & Gudjonsson, 2004; Leo & Ofshe, 1989). Already powerful in its basic form, Kassin and his colleagues found that once innocent suspects are mistakenly identified as guilty, they run the risk of enduring an interview style that is even more grilling than the interview style guilty suspects are subjected to. That is, interrogators who do not believe the innocent suspects' denials are inclined to double their efforts to elicit a confession (Kassin, Goldstein, & Savitsky, 2003).

Perhaps Inbau *et al.* (2001) are aware of the powerful effect that their interrogation technique has on suspects, which may be why they recommend only subjecting suspects to this type of interrogation "...whose guilt, in the *opinion* of the investigator, seems definite or reasonably certain" (italics by Inbau *et al.*, 2001, p. 209). The notion of guilt could come from factual evidence, such as fingerprints, DNA sample, and statement from a reliable witness. However, when such evidence is unavailable, it will most likely come from the demeanour displayed by the suspect during the pre-interrogation interview (Hartwig, 2005). Although Inbau *et al.* (2001, p. 78) report that "The successful interrogator must possess a great deal of inner confidence in his ability to detect truth or deception...", this chapter has demonstrated that we should be sceptical about the average interrogator's ability to detect truths and lies. I believe that judging guilt or innocence from demeanour will lead to many innocent suspects being subjected to powerful interrogation techniques. Innocent suspects who are introverted or socially anxious may run a heightened risk of being interrogated, because their demeanour makes a suspicious impression regardless of whether they are guilty or innocent. (Tom Sawyer, whose case was described above, was a socially anxious man, Leo & Ofshe, 1998.) Also, non-Caucasian suspects may be particularly at risk when interviewed by Caucasian interrogators, because their demeanour also evokes suspicion.

Several veracity assessment tools have been developed for professional use in an attempt to improve people's ability to discriminate between truth tellers and liars. These tools will be discussed in the next seven chapters, starting with a lie detection interview protocol described in the Inbau *et al.* (2001) manual: the Behaviour Analysis Interview.

Appendix 6.1 Accuracy rates of laypersons judging strangers[1]

	Accuracy rate (%)		
	Truth	Lie	Total
Anderson, DePaulo, & Ansfield (2002)			51
C. F. Bond & Fahey (1987)	49	50	49
C. F. Bond, Kahler, & Paolicelli (1995)			63
C. F Bond, Omar, Mahmoud, & Bonser (1990)[2]			56
C. F. Bond *et al.* (1992, experiment 1)			50
C. F. Bond *et al.* (1992, experiment 2)			55
C. F. Bond *et al.* (1992, experiment 3)			52
C. F. Bond *et al.* (2004)	56	47	52
G. D. Bond *et al.* (2004)	74	44	59
G. D. Bond, Malloy *et al.* (2005)	80	27	54
Brandt, Miller, & Hocking (1980a)			38
Brandt, Miller, & Hocking (1980b)			42
Brandt, Miller, & Hocking (1982)			31
Davis, Markus, & Walters (2006)[3]	62	55	58
DePaulo & Pfeifer (1986)			54
deTurck (1991)			54
deTurck, Feeley, & Roman (1997)			57
deTurck, Harszlak, Bodhorn, & Texter (1990)			64
deTurck & Miller (1990)			53
Edelstein, Luten, Ekman, & Goodman (2006)	59	41	50
Ekman & O'Sullivan (1991)			53
Ekman, O'Sullivan, & Frank (1999)			58
Etcoff, Ekman, Magee, & Frank (2000)			47
Feeley & deTurck (1995)	72	54	63
Feeley & deTurck (1997)			49
Feeley, deTurck, & Young (1995)			56
Feeley & Young (2000)[4]	54	46	50
Fiedler & Walka (1993)	55	61	53
Fleming & Darley (1991)			42
Frank & Ekman (1997, experiment 1)[5]			58
Frank & Ekman (1997, experiment 2)			60
Frank, Feeley, Paolantonio, & Servoss (2004)	59	53	56
Garrido, Masip, & Herrero (2004)	51	66	59
Granhag & Strömwall (2000a)	69	44	57
Granhag & Strömwall (2001c)[6]	56	59	57
Hartwig *et al.* (2004b)	50	65	58
A. K. Johnson *et al.* (2004)[7]	60	50	55
A. K. Johnson *et al.* (2005)			55
Kassin & Fong (1999)			56
Kassin, Meissner, & Norwick (2005)	63	54	59
Landström, Granhag, & Hartwig (2005)			50
Lane & DePaulo (1999)			59
Levine & McCornack (2001, study 1)			54
Levine & McCornack (2001, study 2)			59
Levine & McCornack (2001, study 3)			56

(Continued)

Appendix 6.1 (*Continued*)

	Accuracy rate (%)		
	Truth	Lie	Total
Levine, Park, & McCornack (1999, study 2)	70	38	54
Levine, Park, & McCornack (1999, study 3)	69	27	48
Levine *et al.* (2005, study 1)			52
Levine *et al.* (2005, study 2)			53
Levine *et al.* (2005, study 4)	69	32	50
Levine, Kim, Park, & Hughes (2006)	67	34	50
Littlepage & Pineault (1985)[8]	81	30	56
Malcolm & Keenan (2005)	73	44	59
Masip, Garrido, & Herrero (2006)	59	49	54
Mei-tai Fan, Wagner, & Manstead (1985)			58
Millar & Millar (1995, study 1)[9]			50
Millar & Millar (1995, study 2)[9]			50
Millar & Millar (1997a)			49
Miller *et al.* (1981, study 1)[10]			55
Miller *et al.* (1981, study 2)[11]			56
Miller, deTurck, & Kalbfleisch (1983)[12]			51
Newman, Pennebaker, Berry, & Richards (2003)	74	30	52
O'Sullivan (2003)[13]	62	49	55
O'Sullivan, Ekman, & Friesen (1988)	69	48	54
Porter, Campbell, Stapleton, & Birt (2002)			58
Porter, Woodworth, McCabe, & Peace (2007)			53
Santarcangelo, Cribbie, & Ebesu (2004)	70	59	65
Schul, Mayo, Burnstein, & Yahalom (2007)			59
Smith, Archer, & Constanzo (1991)			60
Stiff & Miller (1986)	67	41	54
Strömwall, Granhag, & Jonsson (2003)[14]	70	45	58
Vrij & Baxter (1999)[15]	52	54	53
Vrij & Graham (1997)			42
Vrij, Akehurst, Brown, & Mann (2006)	58	63	61
Vrij, Harden, Terry, Edward, & Bull (2001, study 1)[16]	57	53	55
Vrij, Harden, Terry, Edward, & Bull (2001, study 2)[17]	53	55	54
Wan Cheng & Broadhurst (2005)[18]	67	70	68
Zuckerman, Koestner & Alton (1984)			62
Zuckerman, Koestner, & Colella (1985)[19]			58
Total accuracy scores	**63.41**	**48.15**	**54.27**

Notes: [1]Only control conditions are included; [2]Jordanian and American observers together; [3]The observers watched and listened to videotapes of criminal confessions. This is the only study in Appendix 6.1 using real-life material; [4]Low and high cognitive capacity conditions combined; [5]Scores of crime video and opinion video combined; [6]Single condition; [7]Fake bad and fake good conditions combined; [8]Spontaneous and planned conversations combined; [9]Voice and video information conditions combined; [10]Audio and visual condition, factual and emotional statements combined; [11]Live condition; [12]All conditions combined; [13]Opinion and crime videos combined; [14]Single condition, 1st and 2nd judgements combined; [15]Denials and elaborations combined; [16]Low and high stake lies combined; [17]Easy and difficult lies combined; [18]Cantonese and English conditions combined; [19]Face only, speech only and face plus speech conditions combined

CHAPTER 7

The Behaviour Analysis Interview

This is the first of seven chapters discussing veracity assessment tools used by professional lie catchers and scientists. In Chapters 8 to 10 I introduce tools that are used to analyse speech content. For this purpose the oral statements of senders need to be audiotaped and transcribed, or senders are instructed to write down their statements themselves. The analyses take place on the basis of this written material. I then discuss in Chapters 11 to 13 tools that examine senders' physiological responses, such as skin response, breathing rate, and blood pressure. To measure those responses examinees are wired up to a machine called the "polygraph". A relatively recent development in lie detection is to measure people's brain activity when they are telling the truth or lying. For this purpose electroencephalograms (EEGs) or brain scans (functional magnetic resonance imaging fMRI tests) are carried out.

Many of these tools share a common theme in that they are laborious. Statements need to be audiotaped and transcribed, or examinees need to be wired up to a polygraph, wear a cap containing electrodes (often required for recording EEGs), or be placed in an fMRI scanner. This makes on-the-spot judgements difficult. An exemption is the Behaviour Analysis Interview (BAI), which does not require transcribing or equipment. The BAI stands out from the other tools for another reason: it is the only professional tool that examines behavioural cues to deception.

This chapter deals with the BAI. I discuss the background and rationale of the BAI, followed by a description of the BAI protocol. I then discuss whether it works by reviewing the available BAI research. This chapter reveals that the theoretical principles behind the BAI can be challenged, and that the research findings raise doubts about whether the BAI is an efficient lie detection tool.[1]

BACKGROUND AND PURPOSE OF THE BAI

The BAI is a detection of deception tool developed by John E. Reid and Associates. The groundwork for the BAI was carried out by Frank Horvath who conducted a field study in which he examined the verbal and nonverbal responses of examinees when answering a set of questions prior to their polygraph examinations (Horvath, 1973). The BAI protocol is a modification of the set of questions tested by Horvath. The BAI is taught by John E. Reid and Associates as part of a three-day package that includes other activities such as training in "The Reid Nine Steps of Interrogation" already referred to in Chapter 6. The BAI method, as well as The Reid Nine Steps of Interrogation and other activities, are all described in detail in the manual written by Inbau, Reid, Buckley, and Jayne, *Criminal Interrogation and Confessions*. The first edition appeared in 1962 and the latest, fourth, edition was published in 2001. On their website http://www.reid.com/training_programs/r_interview.html the Reid group report that participants for their training programme "come from both the private sector (retail, finance, health care, manufacturing, etc.) and the public sector, including all levels of law enforcement and Government; from every US State and Canadian Province, as well as countries in Europe, Asia and the Middle East". They further report that more than 300,000 professionals in the law enforcement

[1] I discuss the BAI as it is described in Chapter 11 of Inbau, Reid, Buckley, & Jayne's (2001) manual. After we published an experiment about the BAI (Vrij, Mann, & Fisher, 2006a), Joseph Buckley, President of John E. Reid and Associates and co-author of the Inbau *et al.* manual, responded. First, he wrote an article together with his collaborators (Horvath, Blair, & Buckley, in press). Intriguingly, there are differences in how the BAI is described in that article compared to how it is discussed in the Inbau *et al.* manual. The rationale of the BAI as presented in the article differs from the rationale given in Chapter 11 of the manual. Moreover, the article makes clear that the use of the BAI is restricted and that it can only be applied when certain assumptions are met, whereas Chapter 11 of the manual does not mention any such restrictions. Second, Mr Buckley confided to me in a letter that our article about the BAI contained several errors. Again, the description of the BAI in the letter differed in several aspects from the description of the BAI in Chapter 11 of the Inbau *et al.* manual. I would like to encourage the authors of the Inbau *et al.* manual to update their BAI chapter in the next edition of their manual and to provide a more accurate description of the BAI.

and security fields have attended their three-day programme since it was first offered in 1974. The BAI technique is believed to be one of the two most commonly taught questioning methods in the US (Frank Horvath, 2006, personal communication), and a survey amongst Texas law enforcement officers supports this view (Colwell, Miller, Lyons, & Miller, 2006).

The BAI could be an important first step in police interviewing. The BAI protocol is used in a pre-interrogation setting to shed light on the possible guilt or innocence of the suspect. In other words, investigators could use the BAI for screening purposes to determine the worthiness of further interrogating a suspect. The BAI could also be part of the pre-interview phase of a polygraph examination and could then, together with the polygraph examination, determine the judgement of guilt or innocence of the polygraph examinee (Horvath, 1973).

THE BAI PROTOCOL AND ITS UNDERLYING ASSUMPTIONS

Inbau *et al.* (2001, p. 173) describe the core of the BAI as "the asking of behaviour-provoking questions that are specifically designed to evoke behavioural responses". The core, as well as the term Behaviour Analysis Interview, does not adequately reflect the protocol, because it looks at nonverbal *and* verbal responses of interviewees. The BAI protocol includes asking an open-ended question that invites suspects to describe their activities during a specific period of time (e.g., "What did you do between 3pm and 4pm?"), which is then followed by a series of standardised questions. It is thought that truth tellers and liars will respond differently to these standardised questions. The BAI protocol, as described in the Inbau *et al.* manual in Chapter 11, consists of 15 questions. In his Masters thesis, Blair (1998) added an additional question (the "bait question") to the list. The bait question also appears in Inbau *et al.*'s manual, albeit not as part of the BAI. Instead, it is discussed in the chapter "The use of specialized questions techniques" (Chapter 12). Following Blair, I will discuss the bait question as part of the BAI protocol. An overview of the 16 BAI questions is given in Table 7.1. In this table, I have tailored the questions towards the theft of money from a wallet. For other types of crimes, the questions would need to be slightly rephrased. The emboldened labels are the labels that Inbau *et al.* (2001) gave to these questions, and I use their labels throughout this chapter.

Regarding the *nonverbal responses*, Inbau *et al.* (2001) report that liars feel less comfortable than truth tellers in an investigative

Table 7.1 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, **history/you**. 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?
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 the money? 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?[1]

Source: derived from Inbau, Reid, Buckley, and Jayne (2001)
Note: [1]This bait question is mentioned by Blair (1998) but not by Inbau *et al.* (2001) in their chapter on the BAI or by Horvath, Jayne, and Buckley (1994)

interview situation. As a result, guilty suspects are more likely to cross their legs, shift about in their chair, and perform grooming behaviours while answering the history/you 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. Whilst 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.

Truth tellers and liars are expected to give different *verbal responses* to the 16 questions, because, according to Inbau *et al.* (2001), they have different attitudes towards the investigation. Liars are thought to be less helpful in the interview and do not show an 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 innocent (i.e., truth telling) suspects, guilty (i.e., lying) suspects are thought to be more evasive about the purpose of the interview (e.g., "I have no idea what the purpose of the interview is", question 1); less immediate in their denial of having committed the crime (e.g., "That's where I work, why would I do something like that?", question 2); more likely to deny any knowledge of whom 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 to be innocent (question 5, because guilty suspects prefer to surround themselves with other possible suspects).

Guilty suspects are also more likely to suggest that no crime has actually taken place (e.g., "Perhaps the money got lost?", question 6), and less likely to admit that they had the opportunity to commit the crime (question 7). In addition, guilty suspects are more likely to voice negative feelings about the fact that they are being interviewed (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 thought to be more likely to answer in the third person (e.g., "That's against the law", or "It's wrong"), whereas innocent suspects are more likely to answer in the first person (e.g., "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 that they are being interviewed (question 15).[2]

I believe that the principles underlying the BAI protocol are questionable. The premise underlying the expected differences in nonverbal

[2] Investigators who use the BAI protocol acknowledge that not every response to a BAI question will consistently match the descriptions presented for guilty and innocent suspects. Consequently, investigators should evaluate the responses to the entire BAI interview as a whole rather than to the 16 questions individually.

responses between truth tellers and liars, that is, liars feel less comfortable than truth tellers in an investigative interview, is not universally accepted by the scientific community. For instance, as I explained in Chapter 3, in situations where the consequences of being disbelieved are severe, both liars *and truth tellers* will be concerned about not being believed (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003).

In alignment with this point, but refuting Inbau *et al.*'s assumption, are the results of research examining nonverbal cues to deception. In that research, discussed in detail in Chapter 3, it was 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. This was also found in research where the nonverbal responses of suspects in police interviews were examined (Mann, Vrij, & Bull, 2002). The decrease in movements could be the result of liars having to think harder than truth tellers. If people are engaged in cognitively demanding tasks, their overall animation is likely to reduce (Chapter 3). Alternatively, liars typically experience a greater sense of awareness and deliberateness in their performance because they take their credibility less for granted than truth tellers. Although truth tellers are also keen to be seen as truthful, they typically do not think that this will require any special effort or attention. As a result, liars are more inclined than truth tellers to refrain from exhibiting excessive movements that could be construed as nervous or suspicious (Chapter 3).

This latter impression management explanation (liars put more effort into making a convincing impression than truth tellers) also raises doubts about the view that liars will be less helpful than truth tellers. This was one of the premises underlying the expected differences in verbal responses to the BAI questions between truth tellers and liars. Both truth tellers and liars would probably believe that being helpful makes a positive impression, but perhaps liars will be even keener than truth tellers to make such an impression, because they will take their credibility less for granted.

RESEARCH RELATED TO THE BEHAVIOUR ANALYSIS INTERVIEW

In the opening paragraph of their chapter about the BAI, Inbau *et al.* (2001, p. 173) state that "research has demonstrated that innocent subjects tend to respond differently to these specialized questions than do deceptive subjects". They do not elaborate on this statement, but in an earlier chapter they refer to a field study conducted by Horvath *et al.* (1994) in which the BAI protocol was empirically tested. Horvath *et al.*'s (1994) field study included 60 videotaped BAIs regarding suspected theft that were administered by five interviewers. Those videotapes were rated by four evaluators who were trained and experienced in behaviour analysis interviewing. Those four evaluators classified 78% of the truthful and 66% of the deceptive suspects correctly, but also rated the outcomes for several other suspects as "inconclusive" (that is, neither a truth or lie decision was made). When those inconclusive decisions were disregarded and only the judgements were considered where the BAI evaluators assessed a suspect as being truthful or deceptive, the four evaluators achieved a truth accuracy rate of 91%, a lie accuracy rate of 80%, and a total accuracy rate of 86%.

Although these accuracy rates appear impressive, the study has two important limitations. First, only five interviewers administered the BAIs and these interviews were rated by only four evaluators. These samples are simply too small to obtain insight into the skills of the average BAI interviewer. All four evaluators were highly trained and experienced. As Horvath *et al.* (1994, p. 804) acknowledge, it is unknown whether others with lesser or different qualifications would have achieved equally high accuracy rates in this study. The five interviewers were also experienced in administering BAIs (they had between 8 and 14 years of experience in conducting BAIs), and it is unknown whether the four evaluators would obtain the same accuracy rates if they were to rate BAIs administered by lesser experienced BAI interviewers (Horvath *et al.*, 1994). Second, the ground truth in the study was unclear: it could not be established with certainty that the innocent suspects were truly innocent and the guilty suspects were truly guilty. In that respect this study is far from unique. I already mentioned in Chapter 3 that many field studies suffer from a poor ground truth. 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 for reasons that I will discuss in detail in Chapters 8 and 11. Using a systematic factual analysis is perhaps even more problematic, because facts such as biographical information, opportunity, or motivation do no provide any kind of evidence about the actual innocence or guilt 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). They further report that in only two of the 60 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).

I am sympathetic to the problems faced by Frank Horvath and his team. Establishing ground truth is difficult and very time consuming. In our own field study where we examined behavioural cues to deception displayed by suspects during police interviews, it took the lead researcher, Samantha Mann, more than a year to establish the ground truth, and our study contained only 16 cases (Mann *et al.*, 2002). In our study, establishing ground truth included firstly talking to the detectives who were involved in the investigation, and then going through the often extensive case files related to the crime. Our sample was small because in many cases the ground truth could not be established.

Establishing the ground truth, however, is crucial in field deception research. Only when ground truth is established can researchers conclude with certainty that all the alleged truth tellers in their study were actually telling the truth and that all the alleged liars were actually lying. Certainty about the truth status of all the examinees in a study is essential for drawing any conclusions or for being able to trust the outcomes. To make a comparison, how much faith would a reader have in a study about gender differences if the authors of that study acknowledged that they were uncertain of the gender of their participants? I would not trust those findings and would not rely on the conclusions of the study.

Fortunately, several other studies have been published that have tested the BAI principles in which the ground truth was established. These include studies that examined whether paying attention to the signs of discomfort (legs crossing, gaze aversion, etc.) outlined by Inbau *et al.* would improve lie detection, and studies measuring the cooperativeness of truth tellers and liars. Some, but not all of them, were laboratory-based experiments where undergraduate students were suspects who told the truth or lied for the sake of the experiment.

In one laboratory study, one group of observers were trained to look for the cues of discomfort that Inbau *et al.* (2001) claim to be diagnostic cues to deceit and compared to a group of observers who were not trained (Kassin & Fong, 1999). The observers (undergraduate students) then saw interviews with suspects (undergraduate students) who were either innocent or guilty of a mock crime, and were asked to indicate the truth status (guilty or innocent) of each suspect. The trained observers were less accurate in the lie detection test (46% total accuracy rate) than those who received no training (56% accuracy rate).

This finding, paying attention to Inbau *et al.* cues of discomfort *impairs* lie detection performance, was supported by our field study where police officers judged the veracity of statements made by murder, rape, and arson suspects who told the truth and lied during their real-life (videotaped) police interviews (Mann, Vrij, & Bull, 2004). As already discussed in Chapter 6, we found a negative relationship between officers reportedly attending to the Inbau *et al.* cues (averting gaze, shifting posture, making self-adaptors, etc.) and accuracy in the lie detection task. That is, the more the officers endorsed Inbau *et al.*'s view on cues to deception, the worse they became at distinguishing between truths and lies.

In another laboratory experiment, truth tellers actually took part in an event whereas liars did not take part but were given details about the event and were asked to pretend that they did take part in it (Vrij, 2005a). The truth tellers and liars were subsequently interviewed by a police officer about the event and were asked to recall the event in detail. After they described what had happened, the interviewer challenged the veracity of truth tellers' and liars' accounts, and asked them to report again what had happened. More truth tellers than liars refused to do so, indicating that the truth tellers were *less* helpful than the liars. This contradicts the Inbau *et al.* assumptions.

In a final laboratory experiment we tested the behaviour-provoking questions part of the BAI protocol in our laboratory (Vrij, Mann, & Fisher, 2006a). The truth tellers (undergraduate students) participated in a staged event in which they played a game of Connect 4 with a confederate who posed as another participant. (Connect 4 is a popular two-player game where players drop counters into a slotted grid to achieve, and simultaneously prevent their opponent from achieving, four of their counters in a row.) During the game they were interrupted twice, first by a second confederate who came in to wipe a blackboard and later by a third confederate who entered looking for his or her wallet. Upon finding the wallet, this latter confederate then claimed that a £10 note had gone missing from it. The participant was then told that s/he would be interviewed about the missing money.

The liars did not participate in this staged event. Instead, they were asked to take the £10 from the wallet, but deny having taken this money in a subsequent interview. They were instructed to tell the interviewer that they played a game of Connect 4 like the truth tellers had. The liars were then presented with a sheet containing the information about the staged event that the truth tellers had participated in.

To motivate the truth tellers and liars to perform well in this study, they were told, just before the interview started, that if they convinced the interviewer that they did not take the money, they would receive £15 for participating in this study. If they did not convince the interviewer, they would instead have to write a statement about what actually occurred.

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 BAI questions 4 and 7); Confederate 1 left the room for a while leaving the participant alone in the room, which 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, therefore Confederate 2 could not have taken the money (see question 5).

All participants were interviewed by the same uniformed, male, British police officer. The interviewer was blind to each participant's condition (truth telling or lying). Participants were asked first to report in as much detail as possible what happened when they played Connect 4. After this recall, they were asked 15 behaviour analysis interview questions listed in Table 7.1 (all questions minus the "tell loved ones" question which is not appropriate for testing in a laboratory experiment). All interviews were audiotaped, videotaped, and transcribed. The verbal part of the BAI coding was based on the transcribed responses, and the nonverbal part of the coding was based on the videotapes.

When we analysed the data for each question separately we found little support for Inbau *et al.*'s assumptions. The only question that obtained support for Inbau *et al.*'s predictions was the bait question (question 16). Truth tellers expressed more confidence in being exonerated than liars. Other findings that revealed differences between truth tellers and liars contradicted Inbau *et al.*'s assumptions. Truth tellers were more evasive while answering the purpose question (question 1), and were less likely to name someone who they felt certain did not take the money (vouch question, question 5). Truth tellers were also more likely to cross their legs and were more likely to shift posture than liars. When we, as recommended by Inbau *et al.* (p. 127), calculated a total BAI score (the responses to all the questions combined), we found that truth tellers and liars did respond differently from each other. However, our findings were the total opposite of what Inbau *et al.* predicted. Truth tellers appeared *less* helpful and showed *more* signs of discomfort than liars. In a follow-up study, we showed 68 British police officers a selection of these BAI interviews and asked them to indicate after each interview whether the suspect was telling the truth or lying

(Vrij, Mann, Kristen, & Fisher, 2007). The total accuracy was 51%. This percentage did not differ from, and was almost identical to, the score that could be expected by simply flipping a coin (50%).[3]

EVALUATION OF THE BAI TECHNIQUE

The rationale behind the BAI is that truth tellers and liars will respond differently to the list of behaviour-provoking questions because liars are thought to feel less comfortable than truth tellers, to be less helpful than truth tellers, and to not show the appropriate level of concern about being a suspect. The available laboratory studies and field studies where the ground truth was established do not support this rationale, but show the opposite: truth tellers display more signs of discomfort and appear less helpful than liars.

BAI supporters probably will point out that many of the findings presented in this chapter were obtained in laboratory experiments and will challenge what these studies say about real life. To them I would like to emphasise that the findings of these laboratory studies resembled the findings of our field studies reported in this chapter (Mann *et al.*, 2002, 2004). Further, if they do not wish to accept the findings presented in this chapter, I would encourage them to carry out a proper scientific test of the BAI themselves. I do not condone the exposing of hundreds of thousands of people to a BAI training programme when the benefits of the BAI protocol have not been established, and believe this to be poor practice (see also Blair & Kooi, 2004).

The situation is worrying for innocent suspects who are submitted to the BAI protocol. Investigators may base their impression about the suspect's guilt on the outcomes of a BAI interview. Although there is no guarantee that their impressions will be correct, BAI users may be confident that they are. Police investigators who are reasonably certain of a suspect's guilt may submit a suspect to persuasive interrogation techniques meant to break down the suspect's resistance. As I discussed in more detail in Chapter 6, once innocent suspects are mistakenly identified as guilty, they run the risk of being submitted to an interview style that is even more grilling than the interview style that guilty suspects are subjected to. This may lead to false confessions.

[3] Those British police officers were not trained in the BAI but I don't think this negatively affected their accuracy rates. The truth tellers and liars in the BAI interviews responded verbally and nonverbally in a different way than was predicted in the BAI, and BAI trained observers would not have detected the cues that differentiated truth tellers from liars if they had applied their BAI knowledge.

I do applaud the effort that has gone into the BAI to introduce an interview protocol designed to evoke different nonverbal and verbal responses in truth tellers and liars. The question of how specific interview protocols might benefit nonverbal and verbal lie detection has been ignored to date. I believe that such an approach has potential but it is essential that such interview protocols are grounded in sound theory. I will return to this important issue in Chapter 15. I will then outline some interview protocols that I believe will benefit lie detection.

Box 7.1 Legal considerations of the BAI

Moston (1992) has argued that some of the BAI questions may be legally unacceptable in some countries. For example, the bait question (question 16) implies the possible existence of incriminating evidence, which may be misleading; and the purpose question (question 1) is only relevant if suspects have not yet been told the reason why they are being interviewed. However, legislation in for example England and Wales requires custody officers to inform suspects prior to the interview why they are being questioned.

CHAPTER 8

Statement Validity Assessment

This chapter describes Statement Validity Assessment (SVA), which is probably the most frequently used verbal veracity assessment tool to date. SVA assessments are accepted as evidence in some North American courts (Ruby & Brigham, 1997) and in criminal courts in several West European countries, including Austria, Germany, Sweden, Switzerland, and the Netherlands (Köhnken, 2002, 2004). SVA originates from Sweden (Trankell, 1963) and Germany (Undeutsch, 1967; Arntzen, 1970) and has been designed to determine the credibility of *child* witnesses' testimonies in trials for *sexual offences*. It is not surprising that a technique has been developed to verify whether or not a child has been sexually abused. It is often difficult to determine the facts in an allegation of sexual abuse, since often there is no medical or physical evidence. Frequently the alleged victim and the defendant give contradictory testimony and there may be no independent witnesses to give an objective version of events. This makes the perceived credibility of the alleged victim and defendant important. The alleged victim is in a disadvantageous position if he or she is a child, as adults have a tendency to mistrust statements made by children (Ceci & Bruck, 1995).[1]

[1]It is unclear how many child witness statements about sexual abuse are (partially) inaccurate. American estimates range from 6% to 60% (Craig, 1995). The reported cases of invalid accounts of sexual abuse include pressure from adults and/or peers to give a false statement, misidentification of the alleged perpetrator, and outright fabrications.

After a synopsis of the historical background of SVA, I describe the four phases (stages) it consists of. I then review the available SVA research. The core phase of SVA is Criteria-Based Content Analysis (Berliner & Conte, 1993), a list of 19 criteria thought to be more present in truthful than in false accounts. Most of the research has concentrated on this part of SVA. I discuss how accurate experts are in detecting truths and lies when they apply Criteria-Based Content Analysis (CBCA) and the extent to which two experts who analyse the same statement obtain the same CBCA scores (so-called inter-rater agreement). Research reveals that truths and lies can be detected above the level of chance with the CBCA method, but errors are also made.

I then discuss research involving the Validity Checklist (another phase of SVA), a list of 11 issues other than deception thought to have an impact on CBCA scores. Validity Checklist research includes studies examining the effects of the interviewing style, the interviewee's age, and coaching of the interviewee on CBCA scores. I will show that all three issues impact upon CBCA scores. In this chapter I review the several problems that are associated with applying the Validity Checklist, which include difficulty in identifying and measuring some issues that may influence CBCA scores, and difficulty in determining the exact impact of these issues on CBCA scores. I will also question the justification of some of the issues that are included in the Validity Checklist and argue that some issues that may influence CBCA scores are not included in the Validity Checklist. I finally discuss the implications of the research findings for the use of SVA in criminal investigations. I will argue that SVA evaluations are useful in criminal investigations but are not accurate enough to be admitted as expert scientific evidence in criminal courts.

THE HISTORY OF SVA

Statement analysis, initially less systematic than the current SVA procedure, has been applied by German experts in criminal court cases since the 1950s (Steller & Boychuk, 1992). In 1954 the Supreme Court of West Germany summoned a small number of experts to a hearing. The Court wanted to assess to what extent psychologists could help in determining the credibility of child witnesses' testimonies, particularly in trials for sexual offences. The forensic psychologist Udo Undeutsch reported the case of a 14-year-old alleged victim of rape that he had investigated. The five Justices of the Senate "were impressed by the demonstration and convinced themselves that in assessing the truthfulness of the testimony of a child or juvenile witness an expert psychologist conducting an out-of-court examination has other and better resources than the persons acting as fact finders within the formal atmosphere of a courtroom trial" (Undeutsch, 1989, p. 104).

Subsequently a ruling was made in 1955 by the German Supreme Court that required the use of psychological interviews and assessments of credibility in virtually all contested cases of child sexual abuse. This led to numerous cases in which psychologists were called on as experts. Arntzen (1982) estimated that by 1982 expert testimony had been offered in more than 40,000 cases. In West Germany and Sweden this further resulted in the development of various content criteria to assess the credibility of statements made by alleged victims of sexual abuse. Undeutsch (1967, 1982) was the first to compile a comprehensive list of criteria but others have published similar lists (Arntzen, 1970, 1983; Littmann & Szewczyk, 1983; Trankell, 1972).

With the help of others, Günter Köhnken and Max Steller took statement analysis a step further. They refined the available criteria and integrated them into a formal assessment procedure, which they called Statement Validity Assessment (Köhnken & Steller, 1988; Raskin & Esplin, 1991b; Raskin & Steller, 1989; Raskin & Yuille, 1989; Steller, 1989; Steller & Boychuk, 1992; Steller & Köhnken, 1989; Yuille, 1988b). The procedure developed by Köhnken and Steller is the procedure used to date.

The formal procedure therefore arrived in the late eighties more than 30 years after the German Supreme Court ruling in 1955. It was also after more than 30 years of forensic practice in German courts that the first German study to validate statement analysis was published (Steller, 1989). Obviously, research is needed regarding the validity[2] of a procedure that is used as evidence in criminal courts (Doris, 1994), and this research has arrived since the publication of the SVA protocol in the late eighties. This chapter reviews all the research published in English to date.[3]

SVA is well established in German criminal courts. Prosecutors and defence lawyers rarely challenge the reliability[4] or validity of the test, although they are allowed to do so (Köhnken, 1997, personal communication). Both the prosecution and defence are also allowed to challenge

[2] In psychology "validity" refers to the accuracy of the test, in this case whether SVA actually discriminates between truths and lies.

[3] For other reviews and discussions of the SVA procedure see Horowitz (1991); Köhnken (1996, 2002, 2004); Lamb *et al.* (1997a,b); Lucas & McKenzie (1998); Memon, Vrij, & Bull (2003); Pezdek & Taylor (2000); Porter & Yuille (1995); Ruby & Brigham (1997); Tully (1999); Vrij (2005b); and Zhou, Burgoon, Nunamaker, & Twitchell (2004a).

[4] Reliability refers to consistency between SVA experts in applying the SVA procedure. That is, if two SVA experts assess the same case independently, do they then reach the same outcome?

or discredit SVA evidence, for instance by finding weak spots in the expert witness' reasoning, by cross-examining the expert in court, or by hiring another expert to advise them about the quality of the expertise (Köhnken, 1997, personal communication). I am not aware how well established the use of SVA is in other countries where they apply the method, but it may well be less established than in Germany. Ruby and Brigham (1997, 1998) pointed out that the results of SVA are presented as evidence through expert testimony in some North American courts, but this is considerably less common than in Germany. The main value of SVA in North America seems to lie in its utility for guiding police investigations, and exercising prosecutorial discretion (Raskin & Esplin, 1991b). SVA should be used more widely according to Honts (1994), who argues that its validity has been conclusively demonstrated, and several researchers have pressed for SVA assessments to be accepted as evidence in North American criminal courts (Raskin & Esplin, 1991a, b; Zaparniuk, Yuille, & Taylor, 1995). Others, however, are more sceptical about the procedure (Brigham, 1999; Davies, 2001; Lamb, Sternberg, Esplin, Hershkowitz, Orbach, & Hovav, 1997b; Pezdek & Taylor, 2000; Rassin, 1999; Ruby & Brigham, 1997; Wells & Loftus, 1991). In short, opinion about SVA is divided, and the technique is more established in some countries than in others.

THE FOUR STAGES OF STATEMENT VALIDITY ASSESSMENT

SVA consists of four stages: (i) a case-file analysis to gain insight into the case; (ii) a semi-structured interview to obtain a statement from the interviewee; (iii) a Criteria-Based Content Analysis (CBCA) that systematically assesses the quality of a statement; and (iv) an evaluation of the CBCA outcome via a set of questions (Validity Checklist).

Stage 1: A Case-File Analysis

The SVA procedure starts with the analysis of the case-file (Köhnken, 2004). A case-file should include information about the child witness (e.g., his or her age, cognitive abilities, relationship to the accused person); the nature of the event in question (e.g., whether it was a single event or repeated occurrences); previous statements of the child and other parties involved; and other characteristics such as the time interval between the occurrence of the event and reporting the event, and the relationship between the other parties involved (e.g., whether the parents of the child are involved in a dispute over custody of the child).

The case-file analysis gives the SVA expert the opportunity to formulate hypotheses about what may have happened, and this will influence the activities in the subsequent stages. Suppose that the accused person has admitted to having met the child at a certain location. It is then unnecessary to focus much on the location in the interview in Stage 2, and information about the location provided by the child in the interview should not be included in the CBCA analysis in Stage 3, because this information is not diagnostic of truth telling or lying in this case. And suppose that the parents of the child are involved in a dispute about custody over the child. It may then be that one parent encourages the child to falsely accuse the other parent of having sexually abused the child in an attempt to win the dispute, something the expert should consider in Stage 4. In other words, the case-file analysis may give insight into the issues that are disputed, and the analysis should focus on these disputed elements. It may also give insight into possible suspicious circumstances surrounding the case that the SVA expert should focus on.

Stage 2: The Semi-Structured Interview

The second stage of SVA is a semi-structured interview where the child provides his or her own account of the allegation. Conducting a proper interview is never an easy task, but interviewing young children is particularly difficult, because their descriptions of past events are notably incomplete (Cordon, Pipe, Sayfan, Melinder, & Goodman, 2004; Goodman & Melinder, 2007; Lamb, Sternberg, & Esplin, 2000; Memon, Vrij, & Bull, 2003; Milne & Bull, 1999; Saywitz, 2002). There are several reasons for this, including the fact that young children often lack the cognitive abilities and command of language necessary to offer detailed accounts of events (Davies, 1991, 1994a; Fivush, 2002; Fivush, Haden, & Adam, 1995). Also, young children use simpler and less successful strategies in retrieving memories than do older children and adults, making it more difficult for them to recall information independently (Saywitz, 2002).

Social factors also contribute to incomplete accounts, and this is not just the case for young children. For example, socially anxious people typically feel uncomfortable in the presence of strangers and therefore volunteer less information spontaneously (Saywitz, 2002; Vrij, Akehurst, Soukara, & Bull, 2002). The amount of information provided could also depend on the topic discussed, as some people may feel too embarrassed to talk freely about certain events they have experienced or witnessed (Saywitz, 2002).

For these reasons, interviewers routinely want more information than is initially provided (Kebbel & Milne, 1998), and therefore have to ask further, specific, questions in order to learn more about an event. A commonsense strategy is to ask questions that fit in with the interviewer's understanding of the event, but this can easily result in interviewees agreeing to issues that never happened. The following example of a "conversation" between an adult and a quiet child perhaps looks familiar:

> Adult: "What did you do at school today?" Child does not answer. "Did you go swimming at school today?" Child nods head. "Oh that's nice! Did you enjoy swimming today?" Child nods head again. "Oh good. Did you swim without your armbands?" Child nods again. In fact, the child did not go swimming.

Such a leading questioning style where the child goes along with suggestions made by the adult questioner will not do any harm in many situations. However, it could be damaging if it happens in a criminal investigation. A key element of an appropriate interview is that the child tells his or her own story, without any influence from the interviewer.

The issue of how to properly interview children (and adults) has attracted extensive attention from psychologists. Special interview techniques based upon psychological principles have been designed to obtain as much information as possible from interviewees in a free narrative style without inappropriate prompts or suggestions.[5]

Investigative interviewing research has shown that *open-ended questions* (e.g., "Tell me what happened?") are best for gathering information, as the interviewee has the opportunity to give an open, unrestricted and uninfluenced answer, and has control over the flow of information. This type of question produces responses that are longer and richer in relevant details than responses to other types of questions. Most "Wh-" questions (what, where, when, why, and who) can be classified as open-ended questions. These questions should be asked at the beginning of the interview. *Facilitative responses* by the interviewer such as "okay", "mm", etc. encourage the witness to continue the account. *Focused questions* can be asked later in the interview. These are questions that focus on details or aspects of the event that the witness

[5] Bull, 1992, 1995; Davies, 2004; Davies, Westcott, & Horan, 2000; Goodman & Melinder, 2007; Hershkowitz, 1999; 2001b, 2002; Lamb, Orbach, Sternberg, Esplin, & Hershkowitz, 2002; Lamb, Sternberg, & Esplin, 1994, 1998; Lamb, Sternberg, Orbach, Hershkowitz, & Esplin, 1999; Memon & Bull, 1999; Memon, Vrij, & Bull, 2003; Milne & Bull, 1999, 2006; Raskin & Esplin, 1991b; Sternberg, Lamb, Esplin, Orbach, & Hershkowitz, 2002; Westcott & Brace, 2002; Westcott & Kynan, 2004, 2006.

has described. For example, "Tell me what the man looked like?" is a focused question after an interviewee mentioned a man, but did not describe his appearance.

Leading questions (e.g., "Was the man black?") are problematic as the interviewee may respond by simply repeating the information conveyed in the question, rather than by relying on his or her memory. *Option posing questions* (questions that offer alternatives, e.g., "Was the man black or white?") are better in that respect because they are less leading. However, there is still the chance that the appropriate option may not be incorporated in the alternatives posed by the interviewer. An open-ended question (e.g., "What was the colour of the man's skin?") is a less suggestive question format and therefore preferable. Where leading questions are asked (e.g., "Did he touch you?"), they should be followed by open-ended questions designed to elicit further information about the topic (e.g., "Tell me everything about that").

Few practitioners apply these principles in forensic interviews. For example, in 2002 Helen Westcott and Nicola Brace reported an analysis of 119 interviews with children by police officers and social workers in England and Wales. The analysis revealed that open-ended questions were rarely asked (6%), and that focused questions were most common (47%). Questions that included alternatives occurred more frequently (29%) than did facilitative responses (13%), and some questions were leading (5%) (Sternberg, Lamb, Davies, & Westcott, 2001; Westcott & Brace, 2002). Studies conducted in Israel, Sweden, and the United States provided similar results (Gilstrap, 2004; Westcott & Brace, 2002). Other studies revealed that interviewers find it difficult to maintain and implement the knowledge and skills they should have required during interview training (Westcott & Kynan, 2006; Westcott, Kynan, & Few, 2005).

Stage 3: Criteria-Based Content Analysis

The interviews with the alleged victim are audiotaped and transcribed and the transcripts are used for the third part of SVA: the Criteria-Based Content Analysis (CBCA). The use of transcripts precludes the opportunity to take interviewees' nonverbal behaviour into account when judging the veracity of their statements. Some people believe this is a disadvantage (Landry & Brigham, 1992). The majority of SVA experts, however, believe that nonverbal information will distract the CBCA rater. Indeed, as discussed in earlier chapters, many observers have incorrect, stereotypical beliefs about deceptive behaviour and often make inaccurate judgements when they detect deceit on the basis of someone's behaviour. Videotapes, however, could be useful for a different purpose. For example, a poor interview style and any other possible biasing effects of the interviewer can be identified more easily on videotaped recordings than with transcripts (Honts, 1994; Lamb, Sternberg, & Esplin, 1994; Yuille, 1988b). It is therefore useful to both videotape and audiotape the interview, but the CBCA analysis should take place on the basis of the transcripts only.

Trained CBCA evaluators judge the absence or presence of the 19 criteria that the CBCA list comprises, often on a three-point scale where "0" is assigned if the criterion is absent, "1" if the criterion is present, and "2" if the criterion is strongly present. However, Köhnken (2004) found that using a five-point scale ranging from 0 (criterion is not present) to 4 (criterion is strongly present) is better because such a scale is more sensitive to smaller differences between truthful and fabricated statements. In our own research we use a frequency coding system and rate the number of times each criterion is present in a transcript. The benefits of our method are that we can identify exactly which sentences of the transcript have been given which CBCA code, and, if more than one rater coded the same transcript, where they agree or disagree with each other. Table 8.1 provides the list of 19 CBCA criteria.

Table 8.1 Content criteria for statement analysis

General characteristics
1. Logical structure
2. Unstructured production
3. Quantity of details

Specific contents
4. Contextual embedding
5. Descriptions of interactions
6. Reproduction of conversation
7. Unexpected complications during the incident
8. Unusual details
9. Superfluous details
10. Accurately reported details misunderstood
11. Related external associations
12. Accounts of subjective mental state
13. Attribution of perpetrator's mental state

Motivation-related contents
14. Spontaneous corrections
15. Admitting lack of memory
16. Raising doubts about one's own testimony
17. Self-deprecation
18. Pardoning the perpetrator

Offence-specific elements
19. Details characteristic of the offence

Source: adapted from Steller and Köhnken (1989)

CBCA is based on the hypothesis, originally stated by Undeutsch, that a statement derived from memory of an actual experience differs in content and quality from a statement based on invention or fantasy. This is known as the *Undeutsch hypothesis* (Steller, 1989). CBCA experts believe that each of the 19 CBCA criteria is more likely to occur in truthful than in deceptive statements and it is this more detailed prediction that I have in mind when I refer to the *Undeutsch hypothesis* in this chapter. All CBCA criteria indicate truth. Hence CBCA does not search for lie symptoms, and is therefore not a "verbal lie detector". The absence of a criterion does not necessarily mean that the statement is fabricated (Yuille, 1988b).

The 19 CBCA criteria are clustered into four subcategories. The first three criteria belong to the *general characteristics* category, which refers to the statement as a whole. Köhnken (1996, 1999, 2004) provided a theoretical framework for why differences would occur in Criteria 1 to 3 between truthful and fabricated statements. He suggested a cognitive explanation and argued that the presence of Criteria 1 to 3 is likely to indicate genuine experiences because these criteria are typically too difficult to fabricate.

Criterion 1: Logical Structure. Logical structure is present if the statement is coherent and does not contain logical inconsistencies or contradictions. Logical consistency is not the same as plausibility. Logically consistent statements may sound implausible.

Criterion 2: Unstructured Production. Unstructured production is present if the information is presented in a non-chronological order. Unstructured reproduction occurs in particular when people are upset. For example, highly emotional adult rape victims tend to give their account in unstructured ways (Winkel, Vrij, Koppelaar, & Van der Steen, 1991). Someone may start by explaining the core of the event ("His hands were all over me"), may then go back to the beginning ("He looked OK when I first met him"), and then give information about events that happened later ("He had a cynical smile on his face when he ran away") before going back to an earlier stage of the event ("I should not have been so naive"). This criterion is less useful when someone has already told the story a couple of times or when someone has frequently thought about the event, as this will result in telling a story in a more chronological order.

Criterion 3: Quantity of Details. This criterion is present if the statement is rich in detail and includes specific descriptions of place, time, persons, objects, and events. For example, the following statement would fulfil this criterion: "I used the cash machine on Albert Road near the traffic lights. It was getting dark, and it was drizzling and cold. It was quite busy near the cash machine, at least eight or nine people were standing there. And then, after I took my money and walked away, this guy suddenly appeared, pulled a knife and touched my breasts. I was too scared to do anything and nobody seemed to notice what was going on".

The second category, the *specific contents* category, contains Criteria 4 to 13. These criteria refer to particular sentences in the statement and are meant to reveal the concreteness and vividness of the statement. In fact, they are specifications of the general criterion *quantity of details* (Criterion 3). Similar to Criteria 1 to 3, the presence of Criteria 4 to 13 is likely to indicate genuine experience because these criteria are typically too difficult to fabricate.

Criterion 4: Contextual Embedding. A contextual embedding is present if the events are placed in time and location, and if the actions are connected with other daily activities and/or customs. For example, a victim describes that the crime occurred in a park at lunch-time when the man was walking his dog.

Criterion 5: Descriptions of Interactions. Naturally occurring events typically involve a sequence of actions and reactions, and Criterion 5 addresses this issue. This criterion is fulfilled if the statement contains information that interlinks at least the alleged perpetrator and witness. For example "I said go away but he didn't, he just smiled, and then I started crying" would fulfil this criterion.

Criterion 6: Reproduction of Conversation. Most naturally occurring events involve speech, and Criterion 6 focuses on this aspect. Reproduction of speech is present if parts of the conversation are reported in original form or if the different speakers are recognisable in the reproduced dialogues. This criterion is not satisfied by a report about the content of a dialogue; it is only satisfied when there is a virtual replication of the utterances of at least one person. Thus, "He said: Are you OK, you look so pale?" fulfils this criterion but "He asked how I felt" would not.

Criterion 7: Unexpected Complications during the Incident. This criterion is present if there are elements incorporated in the statement that are somewhat unexpected. For example, if the witness mentions that the alleged perpetrator's car alarm went off at the time of the incident, or that the alleged perpetrator had difficulty with starting his car.

Criterion 8: Unusual Details. Unusual details refer to details of people, objects, or events that are unique, unexpected or surprising but meaningful in the context. Examples are a witness who gives a description of a tattoo on the perpetrator's arm, or a witness who says that the perpetrator had a stutter.

Criterion 9: Superfluous Details. Superfluous details are present if the witness describes details in connection with the allegations that are not essential for the accusation, such as a witness who says that the alleged perpetrator tried to get rid of the cat that entered the bedroom because he (the perpetrator) is allergic to cats.

Criterion 10: Accurately Reported Details Misunderstood. This criterion is fulfilled if the witness mentions details that are beyond his or her comprehension, for example a child that describes the adult's sexual behaviour but attributes it to a sneeze or to pain.[6]

Criterion 11: Related External Associations. A related external association is present if events are reported that are not actually part of the alleged offence but are merely related to the offence, for example, if the interviewee says that the perpetrator talked about sexual affairs with other women.

Criterion 12: Accounts of Subjective Mental State. This criterion is present if the witness describes the development and change of his or her feelings experienced at the time of the incident, such as a witness describing how scared she was during the incident and her relief when it was all over. This criterion also includes reports of thoughts, such as a witness mentioning that she was thinking about how to escape while the event was in progress.

Criterion 13: Attribution of Perpetrator's Mental State. This criterion is present if the witness describes the perpetrator's feelings, thoughts, or motives during the incident: "He was nervous too, his hands were shaking"; and "He thought about the possibility that I would start screaming, because he closed all the windows and played loud music before he started to touch me", and so on.

The third category, the *motivated-related contents* category, includes Criteria 14 to 18. These criteria refer to how the witness presents the statement. According to Köhnken (1996, 1999, 2004) Criteria 14 to 18 are most likely to occur in truthful statements for motivational reasons. A truthful person will not be as concerned with impression management as deceivers (Chapter 3). Compared to truth tellers, deceivers will be keener to try to construct a report that they believe will make a credible impression on others, and will leave out information that, in their view, will damage their image of being a sincere person. As a result, a truthful statement is more likely to contain information that is inconsistent with people's stereotypes of truthfulness. Criteria 14 to 18 are therefore sometimes called "contrary-to-truthfulness-stereotype" criteria (Ruby & Brigham, 1998).

[6] Research suggests that most children under eight years old have hardly any detailed knowledge about sexual behaviour (Gordon, Schroeder, & Abams, 1990; Volbert & Van der Zanden, 1996), and many professionals consider "age-inappropriate" sexual knowledge as an important indicator of sexual abuse (Conte, Sorenson, Fogarty, & Rosa, 1991). However, Jones and McQuiston (1989) argued that there are no age norms for such knowledge.

Criterion 14: Spontaneous Corrections. This criterion is fulfilled if corrections are made or information is added to material previously provided in the statement without having been prompted by the interviewer. "It was about two o'clock, or wait, it must have been later because it was already getting dark" is an example of a correction, and "We were in his car and he drove fast, by the way it was a Volvo, and he drove so fast that he almost failed to stop for a traffic light" is an example of an addition.

Criterion 15: Admitting Lack of Memory. This criterion is present if a witness admits lack of memory by either saying "I don't know" or "I don't remember" or by giving an answer such as "I forgot all about this except for the part when we were in the car." Like Criterion 14, this criterion is only present when the utterance is made without being prompted by the interviewer. Thus, saying "I don't know" or "I can't remember" as an answer to a direct question ("What was the colour of his shirt"?) does not count as admitting lack of memory.

Criterion 16: Raising Doubts About One's Own Testimony. This criterion is present if the witness indicates that part of his or her description sounds odd, implausible, unlikely, etc. (e.g., "You know, this thing is so weird and he seemed to be such a nice man – the whole neighbourhood likes him – that I thought nobody would ever believe me").

Criterion 17: Self-Deprecation. Self-deprecation is present if the witness mentions personally unfavourable, self-incriminating details, for example, "Obviously it was stupid of me to invite him to my home."

Criterion 18: Pardoning the Perpetrator. Pardoning the perpetrator is present if the witness excuses the alleged perpetrator for his or her behaviour or fails to blame the alleged perpetrator, for example, a girl who says she now feels sorry for the defendant possibly facing imprisonment because she does not think it was his intention to hurt her.

Criterion 19: Details Characteristic of the Offence. A single criterion constitutes the final and fourth category *offence-specific elements*. Criterion 19 is present if a witness describes elements that are known by professionals to be typical for this type of crime, but are counterintuitive for the general public. This criterion is likely to be more present in truthful statements because these elements are assumed to be too difficult to fabricate (Köhnken, 1996, 1999, 2004; Marshall & Alison, 2006).

Inter-Rater Reliability Scores

When two people rate a transcript independently from each other, do they then obtain the same CBCA score? This is an important question, because if they do not, it means that the outcome of CBCA coding does not solely depend on the quality of the statement (as it should) but also on the interpretation of the individual CBCA coder. To answer this inter-rater reliability question, I reviewed the available CBCA research. I give a summary of the results here, and the full details are published elsewhere (Vrij, 2005b). For most individual criteria, good, albeit not perfect, inter-rater agreements were obtained in the majority of studies (there were exceptions for Criterion 2, *unstructured production*, and Criterion 14, *spontaneous corrections*). Moreover, the inter-rater agreement scores for the total CBCA scores were typically somewhat higher than those for the individual criteria. These findings suggest that different CBCA coders often agree that a particular sentence in a statement includes a CBCA criterion, but sometimes disagree about which criterion it is. The high but not perfect inter-rater agreement findings suggest that total CBCA scores are reliable but that individual coders sometimes obtain different results. This makes it important that in real life at least two CBCA experts rate a statement rather than just one. Only by asking two CBCA experts to evaluate the same statement independently from each other, can someone determine whether the CBCA score merely reflects the quality of the statement, or, at least in part, the private opinion of the individual CBCA coder. In real-life cases, however, it is common practice for only one expert to examine the transcript.[7]

Stage 4: Evaluation of the CBCA Outcome: The Validity Checklist

A CBCA evaluation is in itself not sufficient to make a decision about the truthfulness of a statement, because CBCA scores may be affected by issues other than the veracity of the statement. For example, the interviewer could have guided the interviewee and could have filled in many gaps in the interviewee's story, or prior to the interview others could have instructed the interviewee about what to say. In those situations even fabricated stories could be rich and detailed. The reverse could also occur. A truthful statement may be of poor quality and lack detail, for example because the interviewee is very young, or verbally unskilled, or too upset to say much, or because the interviewer did not give the interviewee enough opportunity to tell the whole story.

The fact that issues other than veracity influence CBCA scores implies that CBCA is not a standardised test. A standardised test has

[7] See Gödert, Gamer, Rill, & Vossel, (2005), Horowitz (1998), Horowitz, Lamb, Esplin, Boychuk, Krispin, & Reiter-Lavery (1997), and Tully (1998) for more details about inter-rater agreement in CBCA studies and for a debate about the importance of high inter-rater agreement scores.

clear norms that gives it psychological meaning and make interpretation possible (Kline, 1993). An intelligence test is a standardised test. If someone obtains a score of 130, then we know that the person is very intelligent and is more intelligent than someone who obtains a score of 70. Without any norms at all the meaning of a test score is impossible to gauge. Therefore, standardisation of a test is essential.

In an effort to standardise CBCA assessments, a Validity Checklist has been developed which consists of issues thought to be worth examining because they may shed further light on the veracity of the statement. By systematically addressing each of the issues mentioned on the Validity Checklist, the SVA evaluator explores and considers alternative interpretations of the CBCA outcomes. Each time an alternative interpretation is rejected, it strengthens the assumption that the CBCA score accurately reflects the veracity of the statement; each time an alternative interpretation is thought to be plausible, the evaluator should consider whether the CBCA score accurately reflects the veracity of the statement. This evaluation of the CBCA scores is the fourth, and final, stage of SVA. Slightly different versions of the Validity Checklist exist (Raskin & Esplin, 1991b; Steller, 1989; Steller & Boychuk, 1992; Yuille, 1988b). The Validity Checklist presented in Table 8.2 is the list published by Steller and colleagues (Steller, 1989; Steller & Boychuk, 1992). In the present section I limit myself to describing the Validity Checklist. I have concerns about the Validity Checklist which I will discuss later in this chapter in the "Validity Checklist: Reflections" section.

Table 8.2 The Validity Checklist

Psychological characteristics
1. Inappropriateness of language and knowledge
2. Inappropriateness of affect
3. Susceptibility to suggestion

Interview characteristics
4. Suggestive, leading, or coercive questioning
5. Overall inadequacy of the interview

Motivation
6. Questionable motives to report
7. Questionable context of the original disclosure or report
8. Pressures to report falsely

Investigative questions
9. Inconsistency with the laws of nature
10. Inconsistency with other statements
11. Inconsistency with other evidence

Source: adapted from Steller (1989)

The first three issues deal with individual characteristics of the interviewee.

1. **Inappropriateness of Language and Knowledge**. This issue refers to whether the witness' use of language and display of knowledge is beyond the normal capacity of a person of his or her age or beyond the scope of what the witness may have learned from the incident. When this is the case, it may indicate the influence of other people in preparing the statement. For example, to obtain custody a woman may encourage her child to falsely accuse her ex-husband of having abused the child. In an attempt to make a convincing case, the woman may have prepared the statement together with the child and may have coached the child in what to say.
2. **Inappropriateness of Affect**. This issue refers to whether the affect displayed by the witness when being interviewed (usually via nonverbal behaviour) is inappropriate for the witness' alleged experiences. For example, sexual offences are emotionally disturbing and likely to upset victims. One could therefore expect a clear display of emotion from a genuine victim when being interviewed. Absence of these emotions may indicate that the story has been fabricated.
3. **Susceptibility to Suggestion**. This issue refers to whether the witness demonstrates any susceptibility to suggestion during the interview. Some people are more suggestible than others. To use the previous example above regarding questioning a child about whether or not he or she went swimming, some children will actually say correctly that they did not go swimming. Despite considerable controversy concerning the extent to which children and adults are susceptible to suggestion in forensic contexts, three conclusions can be drawn. First, there are age differences in suggestibility, with younger children being more suggestible than older children, and children being more suggestible than adults (Pezdek & Hodge, 1999; Pezdek & Hinz, 2002). Second, there are individual differences in suggestibility (Ceci, Crossman, Scullin, Gilstrap, & Huffman, 2002). Third, the manner in which an interview is conducted can affect suggestibility (as well as the quality and extent of information elicited) (Sternberg, Lamb, Esplin, Orbach, & Hershkowitz, 2002).

Yuille (1988b) and Landry and Brigham (1992) recommend asking the witness a few misleading questions at the end of the interview to assess the witness' susceptibility to suggestion. For example, the interviewer might suggest to the child that the defendant has a fish tank in his living room (which the interviewer knows to be untrue) and observe how the child responds. If the child goes along with these specially devised suggestive questions, this could indicate that he or she is highly suggestible. Obviously these misleading questions should be asked only about peripheral information, and not about central information (e.g., questions about the alleged sexual abuse). The latter will only serve to damage the quality of the statement if a child goes along with the incorrect suggestion, because it means that the child will make a comment about the alleged sexual abuse that the interviewer knows is not true. Also, misleading questions can distort the memory of interviewees, and they may "remember" events that never took place only because the interviewer suggested to them that these events did occur. I discussed this in Chapter 2. Box 8.1 gives a classic example of this phenomenon.

Box 8.1 How the wording of a question can affect memory

Loftus and Palmer (1974) demonstrated how the wording of a question can influence memory. Participants saw a film of a traffic accident and then answered the question: "About how fast were the cars going when they *contacted* each other?" Other participants were asked the same question, except that the verb *contacted* was replaced by either *hit*, *bumped*, *collided*, or *smashed*. Even though all participants saw the same film, the wording of the question affected their answers. The speed estimates (in miles per hour) were 31, 34, 38, 39, and 41, respectively. One week later, the participants were asked whether they had seen broken glass at the incident site. Although no broken glass could be seen, 32% of the participants in the "smashed" condition said that they had, compared to 14% in the "hit" condition. Hence, the wording of the question had influenced the participants' memory of the incident.

The next two issues refer to the interviewer's style or manner whilst conducting the interview.

4. **Suggestive, Leading or Coercive Interviewing**. This issue examines whether the interviewer put suggestions to the interviewee or exerted any kind of pressure, and so on. SVA analyses should not be carried out if the interview was suggestive, leading, or oppressive.
5. **Overall Inadequacy of the Interview**. Factors other than its suggestiveness also determine the quality of an interview. For example, child interviewees often do not realise that they are allowed to say "I don't know" when they do not know the answer

to a question. Instead of admitting lack of memory or knowledge, children (and sometimes adults) have a tendency to answer questions even when they are not sure about the answer. This may well lead to fabrications. It is therefore important that the interviewer makes clear at the beginning of the interview that an "I don't know" answer is acceptable and that the interviewees should say "I don't know" when appropriate (Milne & Bull, 1999; Mulder & Vrij, 1996).

The next three issues explore the motives of the witness in reporting the incident.

6. **Questionable Motives to Report**. This issue refers to whether the witness may have questionable motives to report the incident. It is always possible that someone else encouraged the witness to make a report. It is therefore important to know the relationship between the witness and the accused and be aware of the possible consequences of the accusation for all individuals involved. Relevant in this context is a custody/access dispute or divorce process between the child's parents. As mentioned previously, it is possible that one of the parties in a custody dispute could encourage the child to falsely accuse the other party in an attempt to win that dispute.
7. **Questionable Context of the Original Disclosure or Report**. This issue refers to the origin and history of the statement, particularly the context of the first report. Possible questionable elements in the context of the original disclosure of the accusation should be explored. For example, was the first report voluntary and who (if anyone) asked the witness to report the alleged offence (parents, teacher, psychologist, and so on)?
8. **Pressures to Report Falsely**. This issue deals with the question of whether there are indications that others suggested, coached, pressured, or coerced the witness to make a false report or to exaggerate certain elements in an otherwise truthful report.

The final three issues look at the statement in its relation to the type of crime and previous statements.

9. **Inconsistency with the Laws of Nature**. This issue refers to describing events that are unrealistic or impossible. If a woman claims that she became pregnant as a result of an alleged incestuous relationship, one might check whether that could have been possible given the age of the witness at the time of the relationship.
10. **Inconsistency with Other Statements**. There is often more than one testimony about an event. The witness may have been interviewed about the issue before, or other people may also have been interviewed about the event. This issue refers to the possibility

that major elements in the description of the core of the event are inconsistent with or contradicted by another statement made by the witness or somebody else. It is important to note that this issue is not fulfilled with every inconsistency or contradiction that occurs. Rather, it refers to "*substantial* discrepancies in the description of the *core* of the event" (Köhnken, 2004, p. 54, italics added by me).

11. **Inconsistency with Other Evidence**. This issue refers to the possibility that major elements in the statement are contradicted by reliable physical evidence or other concrete evidence. Again, this issue refers to substantial discrepancies between the available evidence and the description of the core of the event.

A LITERATURE REVIEW OF CBCA STUDIES

Type of Studies

In an attempt to validate CBCA, two types of studies have been conducted. In *field studies* statements made by interviewees in actual cases of alleged sexual abuse have been examined, whereas in *laboratory studies* statements of participants who lied or told the truth for the sake of the experiment have been assessed. As I already explained in Chapter 3, each paradigm has its advantages and the one's strength is the other's weakness.

The statements assessed in field studies have clear forensic relevance as these are statements from real-life cases. The main problem with field studies is establishing the ground truth, that is establishing the truth or falsity of those statements beyond doubt. Good field studies determine ground truth on the basis of criteria that are independent from the witness statement, such as DNA evidence and medical evidence.[8] However, that type of evidence is often not available in real-life cases where CBCA assessments are conducted (Steller & Köhnken, 1989). Therefore, typically criteria such as convictions or

[8] According to Lykken (1988) a scientifically credible field study should fulfil four criteria: (i) the cases selected should be a representative sample; (ii) the statements should be derived from interviews which have been conducted under real-life circumstances; (iii) the statements should be independently scored by at least two evaluators who were blind to case disposition (that is, who were unaware of the ground truth); and (iv) these scores (e.g., CBCA scores) should be compared with the ground truth which has been established by some criterion that is independent of the (CBCA) scores, such as physical evidence.

confessions are used to establish whether a statement is actually true or false.

The problem with these criteria is that they are dependent on the quality of the statement: low quality statements are less likely to result in convictions and confessions than high quality statements. Since CBCA scores also reflect the quality of the statement (poor quality statements result in low CBCA scores and high quality statements lead to high CBCA scores), convictions and confessions cannot be seen as independent from CBCA scores. For example, in some CBCA studies statements were classified as *doubtful* if the judge dismissed the charges (Boychuk, 1991; Esplin, Boychuk, & Raskin, 1988). However, a dismissal does not necessarily mean that the child's statement was false. It might simply be that the abused child was unable to express convincingly to the judge or jury what he or she had experienced. Such an unconvincing but truthful statement would also result in a low CBCA score. To use conviction as confirmation of the ground truth would incorrectly support the hypothesis that statements with low CBCA scores are likely to be false.

Others have used the absence or presence of confessions as the ground truth (Craig, Scheibe, Raskin, Kircher, & Dodd, 1999). If the only evidence against the guilty defendant is the incriminating statement of the child, which is often the case in sexual abuse cases, the perpetrator is unlikely to confess to the crime if the incriminating statement is of poor quality. A poor quality statement will not be perceived as strong evidence and perpetrators often only confess to a crime if they perceive the evidence against them to be strong (Moston, Stephenson, & Williamson, 1992). A poor quality statement will also lead to a low CBCA score and is unlikely to be judged as truthful by CBCA experts. In contrast, if a false incriminating statement is of high quality and judged to be truthful by a CBCA expert, the chances for the innocent defendant to obtain an acquittal decrease dramatically. If there is no chance of avoiding a guilty verdict, it may even for innocent defendants be beneficial to plead guilty in order to obtain a reduced penalty (Steller & Köhnken, 1989). In summary, poor quality (e.g. unconvincing) statements lead to low CBCA scores and decrease the likelihood of obtaining a conviction or confession, and high quality (e.g. convincing) statements lead to high CBCA scores and increase the likelihood of obtaining a conviction of confession, regardless of whether a statement is truthful or fabricated.

As these examples demonstrate, convictions and confessions are likely to yield support for the predicted relationship between CBCA scores and veracity, even in cases where that support is not warranted due to the positive association between convictions and confessions with CBCA scores. This implies that support is likely to be inflated in CBCA field studies where convictions and confessions are used as ground truth.[9]

In laboratory studies there is no difficulty establishing whether a statement is actually true or false, but experimental situations typically differ from real-life situations. Recalling a film someone has just seen (a paradigm sometimes used in laboratory studies) is different from describing an experience of sexual abuse. Therefore, because of this lack of ecological validity, Undeutsch (1984) believes that laboratory studies are of little use in testing the accuracy of SVA analyses.

Clearly, researchers should attempt to make laboratory studies as realistic as possible, and should try to create situations that mimic elements of actual child sexual abuse cases. Steller (1989) has argued that experiences of sexual abuse are characterised by three important elements: (a) personal involvement; (b) negative emotional tone of the event; and (c) extensive loss of control over the situation. The first element could be easily introduced in a laboratory study, the latter two elements are more difficult to introduce for ethical reasons. A popular paradigm in laboratory CBCA research therefore is to invite participants to give an account of a negative event which they have experienced, such as giving a blood sample, being bitten by a dog, etc., or to give a fictitious account of such an event that he or she has not actually experienced. Obviously, the experimenter needs to establish whether the story is actually true or fictitious, for example, by checking with the participants' parents, although this does not always happen in laboratory research.

Box 8.2 Memory for traumatic experiences

Clearly, incidents of sexual abuse are stressful events. It is difficult to investigate how good people are at remembering highly stressful events, as it is not ethically allowable to create stressful events for the purposes of research. To examine people's memory of stressful events, researchers are therefore dependent on real-life cases. However, it is often unclear what has actually happened in real life, which makes it impossible to determine whether someone's recollection of the event is complete and accurate. Peters (1991) investigated what children reported about different stressful

[9] For a discussion about difficulties in establishing whether a statement is true or false in studies of sexual abuse, see Horowitz, Lamb, Esplin, Boychuk, Reiter-Lavery, & Krispin (1996).

events where he did actually know what happened. He investigated children's accounts about a visit to the dentist, an inoculation in a clinic, and a staged theft. His findings revealed that stress sometimes led to an impaired recollection of the event.

The research literature overall, however, provides a different picture: dramatic and traumatic events are typically better remembered than ordinary events (Cordon *et al.*, 2004; Magnussen, *et al.*, 2006). There are also some striking examples documented in the literature of people who were able to give detailed information about stressful events in which they were involved. Pynoos and Eth (1984), for example, interviewed more than 40 children who witnessed the homicide of one of their parents. They found that these children correctly remembered several details of what they had witnessed. Another study revealed that some children who witnessed the rape of their mother could still remember this event very well (Pynoos & Nader, 1988). Jones and Krugman (1986) found that a three-year-old girl who was abducted, sexually abused, and left in a pit by a stranger could give a detailed account of both the event and the stranger. People who were in the concentration camp Erika during the Second World War still had detailed memories about this period more than 40 years later (Wagenaar & Groeneweg, 1990). In the latter study, prisoners were interviewed twice about their stay in the concentration camp, once in the period 1943–1947 and once in the period 1984–1987. This study therefore tells us of the extent to which memories of stressful events can prevail. It was found that camp experiences were generally accurately remembered, although sometimes specific and essential details were forgotten over time. Amongst these were forgetting about being maltreated, and forgetting having been a witness to murder.

Memories go back to early childhood, albeit rarely before three years of age (Cordon *et al.*, 2004; Goodman & Melinder, 2007; Magnussen *et al.*, 2006). One reason for the apparent lack of memory for very early childhood memories is the lack of linguistic capabilities of very young children. Research suggests that children cannot remember events that occurred before they had learned to talk; children appear to remember events once their linguistic capabilities have developed.

Different laboratory studies have used different research paradigms and those that have been used are listed in the tables of this chapter. I have differentiated between: (i) studies where participants actually took part in an event and then were asked to tell the truth or lie about that event (active); (ii) studies in which participants were shown a videotaped event and then asked to tell the truth or lie about that event (video); (iii) studies in which participants watched a staged event and then asked to tell the truth or lie about that event (staged); and (iv) studies in which participants were asked to tell a truthful or fictitious story about a previous negative experience in their life (memory).

As mentioned before, CBCA was developed to evaluate statements from *children* in *alleged sexual abuse* cases. Many authors still describe CBCA as a technique solely developed to evaluate statements made by children in sexual offence trials (Honts, 1994; Horowitz, Lamb, Esplin, Boychuk, Krispin, & Reiter-Lavery, 1997). Others, however, advocate the additional use of the technique to evaluate the testimonies of adults who talk about issues other than sexual abuse (Köhnken, 2004; Köhnken, Schimossek, Aschermann, & Höfer, 1995; Porter & Yuille, 1996; Ruby & Brigham, 1997; Steller & Köhnken, 1989). These authors have pointed out that the underlying Undeutsch hypothesis is restricted neither to children, witnesses and victims, nor to sexual abuse. To shed light on this issue, I also indicated in the tables whether the statements were derived from children or adults, and whether they were victims/witnesses or suspects. I have labelled participants who discussed a negative life event they had experienced "victims". Finally, I have made a distinction between field and laboratory studies.[10]

In this chapter I review all CBCA/SVA articles and book chapters published in English that I am aware of, which in total is more than 50 studies. Several German researchers have conducted CBCA/SVA research, which is not surprising given the importance of SVA in German criminal investigations. However, the vast majority of those studies are written in German, and therefore not accessible to many readers. Therefore I have left those German publications out of the review. However, I have included some CBCA/SVA papers that were presented at conferences rather than published in journals. I am typically reluctant to discuss conference papers, as they have not necessarily been reviewed by experts in the field, and the quality of the reported research has therefore not been thoroughly checked. I nevertheless have included those conference papers that have been influential in the CBCA/SVA

[10] I believe that the atheoretical nature of CBCA is the reason why CBCA experts disagree about when it can be used. Indeed, Köhnken presented a theoretical underpinning by suggesting that cognitive and motivation factors are responsible for the differences in CBCA scores between truth tellers and liars. However, these are post-hoc explanations and were introduced well after the introduction of the CBCA method. It is also unknown whether his theoretical assumptions are shared by other CBCA experts.

debate, and have been discussed by others in the SVA literature (e.g., in Bradford, 1994; Bybee & Mowbray, 1993; Davies, 2001; Horowitz, 1991; Köhnken *et al.*, 1995; Steller, 1989). Finally, I have included Boychuk's (1991) unpublished field study. This doctoral dissertation has also been influential and has received extensive coverage in the SVA literature (e.g., in Lamb, Sternberg, Esplin, Hershkowitz, & Orbach, 1997a; Horowitz, 1991; and Ruby & Brigham, 1997).

Because the CBCA criteria indicate truth rather than deceit, I have changed the connotations in this chapter compared to previous chapters. In Appendix 8.1 a ">" means that a criterion was *more* often present in *true* statements than in *false* statements; a "<" means that a criterion was *less* often present in *true* than in *false* accounts; if no difference was found between true and false statements, a "–" is reported. A "*" indicates that the particular criterion was not investigated in that particular study. The Undeutsch hypothesis predicts that all criteria are more likely to be present in truthful statements than in deceptive statements.

Differences between Truth Tellers and Liars in CBCA Scores: Field Studies

The first field study ever presented to examine the impact of veracity on CBCA scores was Esplin *et al.*'s (1988) study. Forty statements of allegedly sexually abused children between the age of 3.5 and 17 years old were analysed. A statement was classified as *confirmed* if the accused made a confession or if there was medical evidence that unequivocally corroborated vaginal and/or anal trauma. A statement was classified as *doubtful* if there was persistent denial by the defendant, and/or judicial dismissal and no prosecution. A trained CBCA evaluator rated the statements. If a criterion was absent in the statement, it received a score of 0; if it was present, it received a score of 1; and if it was strongly present it received a score of 2. Hence, total CBCA scores could range between 0 and 38.

The results were striking. The doubtful cases received on average a CBCA score of 3.6, the confirmed statements received an average score of 24.8. Moreover, the score distributions of the doubtful and confirmed groups did not show a single overlap. The highest score in the doubtful group was 10 (one child received that score, three children obtained a score of 0), whereas the lowest score in the confirmed group was 16 (one child obtained that score, the highest score was 34). By assessing differences between the two groups for each criterion, differences between the doubtful and confirmed groups emerged for 16 out of 19 criteria, all in the expected direction. That is, the criteria were more often present in the confirmed cases than in the doubtful cases, which strongly supports the Undeutsch hypothesis (see Appendix 8.1). Wells and Loftus (1991, p. 169) refer to these findings as "among the most impressive we have ever encountered in a psychological study". Lamers-Winkelman (1995) describes them as "too good to be true".

Esplin *et al.*'s study has been heavily criticised. One problem with it was that only one evaluator scored the transcripts. We therefore cannot know whether the coding was reliable. In addition, Wells and Loftus (1991) pointed out that the differences between the two groups could have been caused by age differences between them. Indeed, the children in the doubtful group were younger (on average 6.9 years) than of those in the confirmed group (on average 9.1 years). Moreover, the doubtful group included eight statements from children who were younger than five years old, whereas the confirmed group only contained one statement from a child of under five years. Thirdly, apart from medical evidence, none of the criteria used to obtain the ground truth were independent case facts. I discussed above the problems associated with using confessions to establish the ground truth (this was one of the criteria used to classify cases as confirmed). In addition, judicial dismissal, absence of prosecution, and persistent denial by the defendant (criteria used to classify cases as doubtful) may have occurred not only where the reports were false (as suggested by the researchers) but also where the children were unconvincing even though they were telling the truth and sexual abuse actually did take place. Or, in Wells and Loftus' (1991, p. 169) own words:

> They might have been unconvincing because these children might have deficiencies in logical reasoning, they might have been frightened to the extent that they could not process peripheral detail, they might have poor verbal skills, and so on. Because they are unconvincing witnesses, prosecutors might be unlikely to press charges (lack of prosecution), judges might feel that conviction is unlikely (judicial dismissal), and defense attorneys might be unlikely to advise their clients to admit to the charges (persistent denial by the accused).

In her subsequent study, Boychuk (1991) analysed the statements of 75 alleged sexually abused children, between the ages of 4 and 16 years old. She addressed some of the criticisms of the study she carried out with Esplin and others. For example, the statements were analysed by three raters who were blind to case disposition (truth or false). She also included in her sample, apart from confirmed and doubtful groups, a third *likely abused* group. The likely abused cases were those without medical evidence but with confessions by the accused or criminal sanctions from a Superior Court. Unfortunately, in all of her analyses, including the analysis presented in Appendix 8.1, Boychuk combined the confirmed group and the likely abused group, which makes the ground truth in this combined group unreliable. By assessing differences between the two remaining groups on each criterion, Boychuk found less significant differences than Esplin and colleagues (see Appendix 8.1), but all 13 differences were in the expected direction. That is, the criteria were more often present in the confirmed cases than in the doubtful cases, which again supports the Undeutsch hypothesis.

CBCA assessments were carried out to assess the veracity of adult rape allegations in a field study published by Parker and Brown (2000). Differences were found on several criteria and all differences were in the expected direction (Appendix 8.1). However, this study also had serious methodological problems. For example, the criteria to establish the actual veracity of the statements, (i) convincing evidence of rape (no information is given as to what this meant), and (ii) corroboration in the legal sense with either a suspect being identified or charged, are either too vague or not independent case facts. Also, only one evaluator examined most of the cases. It is also unclear whether the evaluator was blind to the case facts or if s/he had been given any background information about the cases s/he was asked to assess.

Craig *et al.* (1999) examined 48 statements from children between the ages of 3 to 16 who were alleged victims of sexual abuse. A statement was classified as *confirmed* if the accused made a confession and/or failed a polygraph test. A statement was classified as *highly doubtful* if the child provided a detailed and credible recantation and/or when a polygraph test suggested that the accused was innocent. In other words, this study did not establish independent case facts either. Only 14 criteria were used. If a criterion was absent a score of 0 was given, and if a criterion was present a score of 1 was allocated. Only total CBCA scores were examined and scores could vary between 0 and 14. The average CBCA scores of the confirmed cases (7.2) were slightly higher than the average scores of the doubtful cases (5.7).

Rassin and van der Sleen (2005) examined 27 true and 14 false allegations of sexual offences. Unfortunately, they did not report the age of the alleged victims. An allegation was considered *true* if the allegation resulted in a conviction of the accused, whereas an allegation was considered *false* if the victims themselves were convicted for making a false allegation. The authors concluded that their ground truth was thus established from a judicial perspective. This is true, but, once again, convictions are not independent case facts. They looked at seven CBCA criteria of which two discriminated between truth tellers and liars. Both of those criteria occurred more often in the true allegations than in the false allegations, which is in agreement with the Undeutsch hypothesis.

In a better controlled field study where the ground truth was established, Michael Lamb and his colleagues analysed the statements of 98 alleged victims of child sexual abuse (aged 4 to 12 years) (Lamb, Sternberg, Esplin, Hershkowitz, Orbach, & Hovav, 1997b). They included only cases where there was evidence of actual physical contact between a known accused and the child and an element of corroboration was present. Using these selection criteria meant that many cases needed to be disregarded, as the initial sample consisted of 1,187 interviews. If a criterion was absent in the statement, it received a score of 0; if it was present, it received a score of 1. They found fewer significant differences than, for example, Boychuk or Esplin *et al.*, partly because 14 rather than 19 criteria CBCA criteria were analysed. However, again, all differences were in the expected direction: the criteria were more often present in the plausible group than in the implausible group. They also calculated total CBCA scores, which could vary between 0 and 14. Significantly more criteria were present in the plausible group (6.74) than in the implausible group (4.85). This difference, however, is much smaller than the difference found by Esplin and colleagues.[11]

Summary

In five of the six CBCA field studies the ground truth has not been satisfactorily established. This means that caution is required when interpreting the findings of these five studies because they cannot really be trusted. However, a trend emerges when all of the field studies are taken together. The only reliable field study published to date showed

[11] In a case study published by Orbach and Lamb (1999) the accuracy of a statement provided by a 13-year-old sexual abuse victim could be established with more certainty than in most other studies. Information given by the victim during the interview was compared with an audiotaped record of that incident. The child had told her mother that her grandfather had sexually molested her on several occasions but the mother did not believe the allegations. One day, when the grandfather entered the bathroom while the child was listening to music played on an audiotape recorder, she pressed the record button and recorded the sexually abusive incident that was about to unfold. Orbach and Lamb conducted a CBCA analysis on the statement and found that 10 out of the 14 criteria they assessed were present in the statement. Obviously, the results of a study in which only one (truthful) statement is examined, does not say much about the validity of CBCA. Also, the fact that the child knew that there was audiotaped evidence of the incident might have influenced her statement in an unspecified manner. Nevertheless, the nature and strength of the corroborative evidence makes the study worth mentioning.

several, albeit small, differences between truthful and fabricated statements (Lamb *et al.*, 1997b), and all of these differences were predicted by the Undeutsch hypothesis: the criteria were more often present in truthful than in fabricated statements. The other five, less reliable, studies support this pattern. If differences between truthful and fabricated statements emerged, it was always in the predicted direction, with more CBCA criteria being present in the truthful statements than in the false accounts.

Sometimes no differences between truthful and false statements were found, indicating that CBCA criteria are not the equivalent of Pinocchio's growing nose. However, compared to the results for nonverbal behaviours (Appendix 3.1), an important difference emerges. The findings for nonverbal behaviours reveal erratic patterns. In some studies liars displayed a particular behaviour *more* often than truth tellers; in other studies liars displayed the same behaviour *less* often than truth tellers; and in yet another set of studies no relationship was found between the particular behaviour and deception. The results for CBCA criteria are more consistent. No field study has ever found any CBCA criterion to emerge more often in fabricated statements than in truthful statements.

Differences between Truth Tellers and Liars in CBCA Scores: Laboratory Studies

Compared to most field studies, the laboratory studies typically revealed fewer differences between truth tellers and liars (Appendix 8.1). Almost all differences, however, were in the direction predicted by the Undeutsch hypothesis, and the criteria occurred more frequently in truthful reports than in deceptive reports. Many of the findings in laboratory studies that oppose the Undeutsch hypothesis were obtained in the studies carried out by Landry and Brigham (1992) and Ruby and Brigham (1998), and several explanations for their findings are possible. They used raters who were trained for only 45 minutes in CBCA scoring, and it is doubtful whether people can be effectively trained in using CBCA within such a short period of time. Also, raters coded very short statements, on average 255 words, whereas the CBCA method has been developed for use on longer statements (Raskin & Esplin, 1991b). Finally, in Landry and Brigham's (1992) study, some raters did not read transcripts of the statements (common CBCA procedure) but watched videotaped statements instead. As I already explained, CBCA experts are not in favour of assessing videotaped statements. In other words, these studies deviated considerably from the typical

CBCA procedure in several ways, which may have affected the CBCA evaluations.

Robustness of the Findings

CBCA findings are robust if the Undeutsch hypothesis received support in different research settings. Appendix 8.1 shows that this is the case. The CBCA criteria emerged more frequently in truthful than in deceptive statements regardless of the experimental research paradigm that was used, that is actual involvement, watching a video or staged event, or statements derived from memory. In addition, the Undeutsch hypothesis received support in the analysis of statements from children, adults, witnesses, victims, and suspects. This was even found in experiments where statements of children and adults (Akehurst, Köhnken, &, Höfer, 2001 Vrij, Akehurst, Soukara, & Bull, 2002, 2004a, b) and witnesses/victims and suspects were directly compared (Vrij, Akehurst *et al.*, 2002, 2004b). These findings support those CBCA experts who argue that CBCA assessments are not restricted to evaluating statements of children in sexual abuse cases.

SUPPORT FOR EACH INDIVIDUAL CRITERION AND TOTAL CBCA SCORE

The bottom of Appendix 8.1 shows the support for the Undeutsch hypothesis for each criterion. To measure this support, for each criterion I have calculated; (i) in how many studies the criterion occurred more often in truthful than in false accounts; and (ii) in how many studies the criterion was examined. Dividing (i) by (ii) gives the support score. Thus, if a criterion occurred more often in truthful than in deceptive accounts in five out of the 10 studies in which it was examined, the support score is 5/10 = 50%. This percentage would be low if in the other five studies the criterion occurred more often in false accounts than in truthful accounts. However, as I already mentioned, this rarely happened. If the Undeutsch hypothesis was not supported, this was mostly because no differences were found between true and false accounts. In that light, a 50% support score is considerably more impressive. The bottom of Appendix 8.1 shows that Criterion 3, *quantity of details*, received the most support. The amount of details was calculated in 29 studies (field studies and laboratory studies combined), and in 22 of those studies (76%) truth tellers included significantly more details into their accounts than liars. Moreover, in not a single study did truth tellers include significantly less details into their statements than liars. This is impressive support for Criterion 3.[12]

In fact, the support is so impressive, that it is useful to see why quantity of details was not a diagnostic cue to deceit in some studies, including in our own work (Vrij, Mann, Kristen, & Fisher, 2007). In our experiment liars were requested to pretend to have participated in an event that the truth tellers had actually participated in. We then gave our liars detailed information about the event the truth tellers were involved in. Perhaps such detailed information provides liars with the opportunity to include a lot of detail in their stories. However, other CBCA criteria still revealed differences between truth tellers and liars in our experiment, and different types of specific details such as *contextual embeddings* (Criterion 4), *reproductions of conversations* (Criterion 6), or *unusual details* (Criterion 8) appeared more often in truthful than in deceptive statements. In other words, it was not the *quantity* of the details (Criterion 3) but the *quality* of these details (Criteria 4, 6, and 8) that gave the participants' lies away.

Appendix 8.1 further shows that *unstructured production* (Criterion 2), *contextual embeddings* (Criterion 4), and *reproduction of conversation* (Criterion 6) all received considerable support with support scores of at least 50%. Typically, the so-called motivational criteria, Criteria 14 to 18, received less support than most cognitive criteria (Criteria 1–13 and 19). In fact, Criterion 17, *self-deprecation*, has received no support at all to date, despite having been examined in 10 studies. In two studies a significant difference between truth tellers and liars emerged, and both times the criterion appeared *less* often in the truthful statements.

Some researchers have criticised the motivational criteria. Berliner and Conte (1993) pointed out that some motivational criteria, such as Criterion 15 *admitting lack of memory*, require the witness to exhibit a lack of confidence in their account as evidence for truthfulness. This, they noted, suggests by implication that confidence diminishes the likelihood of truthfulness, which they find disputable. Appendix 8.1 further

[12]Many non-CBCA studies have also focused on differences between truthful and fabricated statements in *quantity of details* (Criterion 3) and *contextual embeddings* (Criterion 4). Generally, these studies provide further evidence for the Undeutsch hypothesis, with truth tellers including more details (e.g., Burgoon, Buller, Ebesu, White, & Rockwell, 1996; Jones & McGraw, 1987; Jones & McQuinston, 1989; Köhnken & Wegener, 1982; Lindsay & Johnson, 1987; Pezdek & Hodge, 1999; Pezdek & Roe, 1997) and more contextual embeddings into their accounts than liars (e.g., Alonso-Quecuty, 1991; Johnson & Foley, 1984; Johnson, Foley, Suengas, & Raye, 1988; Johnson & Raye, 1981). Even more studies examining details and contextual embeddings are discussed in Chapter 9, because these criteria are also part of the Reality Monitoring tool.

shows that several researchers did not examine most of the motivational criteria, mainly because they believe it is difficult to code them reliably (Horowitz, Lamb *et al.*, 1997).

As I already discussed in Chapters 3 and 4, focusing on individual cues is not a good strategy to detect lies because it is the equivalent of searching for Pinocchio's growing nose. This also applies to conducting CBCA evaluations. In 20 studies researchers computed total CBCA scores and compared the scores for truth tellers and liars. The results are impressive. In 16 out of 20 studies (80%) the hypothesis that truth tellers will obtain higher total CBCA scores than liars was supported. Only in one of the 20 studies (5%), did truth tellers obtain lower CBCA scores than liars (Ruby & Brigham, 1998). I already discussed above that the protocol used in Ruby and Brigham's (1998) study differed in several ways from the typical CBCA approach. In that respect, Ruby and Brigham's study is not a fair test of the CBCA method.

CORRECT CLASSIFICATIONS OF TRUTH TELLERS AND LIARS ON THE BASIS OF CBCA SCORES

Are Trained CBCA Evaluators Better than Laypersons?

To answer the question of whether CBCA experts are better at classifying truth tellers and liars than laypersons, it is useful to know how someone becomes a CBCA expert. Unfortunately, little is known about what kind of training is actually required for this. According to Raskin and Esplin (1991b), a two- or three-day workshop is advisable, whereas Köhnken (1999, 2004) recommends a three-week training course. Although it is debateable how much training is needed, it probably needs to be a rather extensive training programme. Making CBCA/SVA assessments is never a straightforward task. The CBCA list is extensive and several criteria are difficult to grasp. After the CBCA coding, one needs to assess the issues listed on the Validity Checklist, and this is often a complicated task (Steller, 1989; Wegener, 1989). It is impossible to require the necessary skills to do this without extensive training, and even a two- or three-day workshop may not be sufficient.

Several researchers have directly compared the truth/lie detection skills of trained CBCA coders with those of laypersons (Akehurst, Bull, Vrij, & Köhnken, 2004; Köhnken, 1987; Landry & Brigham, 1992; Ruby & Brigham, 1998; Santtila, Roppola, Runtti, & Niemi, 2000; Steller, Wellershaus, & Wolf, 1988; Tye, Amato, Honts, Kevitt, & Peters, 1999). However, all of these studies clearly fell short of the minimum two- or three-day workshop requirement to become a CBCA expert. The shortest training session (45 minutes) was given by Landry and

Brigham (1992) and Ruby and Brigham (1998), though at 90 minutes, Steller *et al.*'s (1988) training session was not much longer. Akehurst *et al.*'s (2004) session lasted two hours, whereas Köhnken (1987) and Santtila *et al.* (2000) did not provide information about the length of their training sessions. However, the content of their training sessions resembled that of the other studies mentioned in this paragraph, and may thus have been of similar length. In a typical study, trainees are given a handout with information about CBCA criteria. A CBCA trainer then explains the criteria in more detail and gives some examples. Trainees are then asked to rate one or a few exercise statements, which are subsequently discussed with the trainer.

The results of studies where the trained and untrained coders were compared are mixed. Trained judges were better at distinguishing between truths and lies than lay evaluators in some studies (Landry & Brigham, 1992; Steller *et al.*, 1988; Tye *et al.*, 1999), whereas other studies found no training effect (Ruby & Brigham, 1998; Santtila *et al.*, 2000) or found that training made observers worse at distinguishing between truths and lies (Akehurst *et al.*, 2004; Köhnken, 1987). It would be unfair to discredit the CBCA method on the basis of these findings, given the lack of depth of the training sessions. All we can conclude is that providing observers with short CBCA training programmes has an unpredictable effect on their ability to detect truths and lies.[13]

Different Ways of Calculating Accuracy Rates: Weight of the Criteria and Decision Rules

In CBCA research, accuracy rates – the correct classifications of truth tellers and liars – are calculated in three different ways. In most laboratory studies, CBCA scores are subjected to statistical analyses, typically discriminant analysis. In such analyses the CBCA scores are entered into a computer and are used to classify participants as either truth tellers or liars. This classification is compared with the actual status of the participant, and the accuracy rate equals the percentage of correct classifications.

[13] In the only field study related to being trained in using CBCA, Gumpert, Lindblad, and Grann (2002b) compared expert testimony reports prepared by professionals with a statement analysis background to reports prepared by a more clinically oriented group, often employed within child and adolescent psychiatry. They found that the reports of the statement analysis background group were generally of higher quality (see Gumpert, Lindblad, & Grann, 2002a, for how quality was measured). Unfortunately, this study does not tell us much about the effectiveness of CBCA training. As the authors acknowledge, they did not assess the accuracy of the recommendations made in the reports. Moreover, the groups could have differed from each other in ways besides training.

A second method to calculate accuracy rates is by asking CBCA coders to make their own truth/lie classifications. This is what happens in real life. CBCA coders may either use their total CBCA scores to make these classifications or could look at the individual criteria. Steller and Köhnken (1989) noted that some criteria may be of more value in assessing truthfulness than others. For example, the presence of accurately reported but misunderstood details in a statement (Criterion 10), such as a child who describes the adult's sexual behaviour but attributes it to a sneeze or to pain, is apparently more significant than the child describing where the alleged sexual encounter took place (Criterion 4). However, no guidance is given in the SVA procedure about what weight should be given to each of the different criteria, leaving this to the interpretation of the individual coder.

In only one study were differences in truth/lie classifications between computer analyses and CBCA coders compared (Vrij, Kneller, & Mann, 2000). No differences emerged in overall accuracy. However, the computer analysis was more accurate in detecting lies (accuracy rates were 80% and 60% for computer analysis and CBCA coder, respectively), whereas the CBCA coder was more accurate in detecting truths (accuracy rates were 80% and 53% for CBCA coder and computer analysis, respectively; see Table 8.3). The differences in lie and truth accuracy rates between computer analysis and CBCA coder supports the idea that the CBCA coder's decisions are not necessarily based on the total CBCA scores.

A third method of making veracity judgements is by using general decision rules. This method is applied both in research and real life. For example, Yuille (cited in Horowitz, 1991) uses as a rule that in order to judge a statement as truthful "CBCA Criteria 1–5 should be present plus two others". Raskin (cited in Zaparniuk *et al.*, 1995) considers a statement to be truthful if Criteria 1 to 3 plus any four of the remaining criteria are present. Craig (1995) uses yet another general decision rule. He argued that statements containing more than five criteria are likely to be truthful. However, the use of general decision rules has serious shortcomings. It is arbitrary because different experts use different rules, and no expert provides a solid justification for why they use their particular decision rule. More importantly, the SVA assumption that CBCA scores could depend on issues other than veracity, such as age and interview style, is ignored when such general decision rules are applied. Using general decision rules thus violates the SVA procedure. Unsurprisingly, many CBCA experts, including the two German founders, are strongly opposed to the use of decision rules (Steller & Köhnken, 1989).

Table 8.3 CBCA accuracy rates

					Accuracy rate (%)		
Authors	Age	Event	Status	Assessment	Truth	Lie	Total
Field studies							
Esplin *et al.* (1988)	3–15	field	victim	CBCA coders	100	100	100
Parker & Brown (2000)	adult	field	victim	decision rules	88	92	90
Laboratory studies							
Akehurst *et al.* (2001)	7–11/adult	active	n/a	discriminant	73	67	70
Akehurst *et al.* (2001)	7–11	active	n/a	discriminant			71
Akehurst *et al.* (2001)	adult	active	n/a	discriminant			90
Akehurst *et al.* (2004)	7–11	active	n/a	CBCA coders			63
Erdmann *et al.* (2004)[1]	6–8	memory	victim	discriminant	77	66	71
Granhag *et al.* (2006)	12–13	staged	witness	discriminant			58
Höfer *et al.* (1996)	adult	active	n/a	discriminant	70	73	71
Joffe & Yuille (1992)[2]	6–9	active	n/a	CBCA coders			71
Köhnken *et al.* (1995)	adult	video	witness	discriminant	89	81	85
Landry & Brigham (1992)	adult	memory	victim	CBCA coders	75	35	55
Ruby & Brigham (1998)	adult white	memory	victim	discriminant	72	65	69
Ruby & Brigham (1998)	adult black	memory	victim	discriminant	67	66	67

(Continued)

Table 8.3 *(Continued)*

Authors	Age	Event	Status	Assessment	Accuracy rate (%) Truth	Lie	Total
Santtila *et al.* (2000)	7–14 (total)	memory	victim	regression	69	64	66
Sporer (1997)	adult	memory	victim	discriminant	70	60	65
Steller *et al.* (1988)	6–11	memory	victim	CBCA coders	78	62	72
Strömwall, Bengtsson *et al.* (2004)	10–13	active	witness	discriminant	44	64	54
Tye *et al.* (1999)	6–10	active	witness	discriminant	75	100	89
Vrij, Akehurst *et al.* (2004a)[3]	5–6	active	witn/susp	discriminant	71	64	69
Vrij, Akehurst *et al.* (2004a)	adult	active	witn/susp	discriminant	58	65	62
Vrij, Akehurst *et al.* (2004b)[4]	6–15/adult	active	witn/susp	discriminant	50	69	60
Vrij, Edward *et al.* (2000)	adult	video	witness	discriminant	65	80	73
Vrij, Kneller *et al.* (2000)[5]	adult	video	witness	discriminant	53	80	67
Vrij, Kneller *et al.* (2000)[6]	adult	video	witness	CBCA coders	80	60	70
Yuille (1988a)	6–9	memory	victim	CBCA coders	91	74	83
Zaparniuk *et al.* (1995)[7]	adult	video	witness	decision rule	80	77	78
Total accuracy scores					**70.81**	**71.12**	**70.47**

Explanation of the signs: n/a = participants participated in an activity, but they were neither a victim nor a suspect

Notes: [1]First interviews only; [2]Lightly coached condition only; [3]Uninformed participants only; [4]Lightly coached condition only; [5]Uninformed liars only; [6]Uninformed liars only; [7]Accuracy rates apply for the decision rule 'Presence of criteria 1–5, and any two of the remaining criteria'

Accuracy Rates in Field and Laboratory Studies

There are only two field studies that reported accuracy rates or where I could calculate those rates myself on the basis of the reported data (Esplin *et al.*, 1998; Parker & Brown, 2000). Both studies showed very high accuracy rates (Table 8.3). However, as I discussed earlier, both studies had serious methodological flaws, and I therefore prefer to disregard these results.

All other studies that reported accuracy rates were laboratory studies. In these studies, the overall accuracy rates varied from 54% to 90%. One of the studies with a relatively low accuracy rate was Landry and Brigham's (1992) study. I have already given several reasons to explain their exceptional findings. In addition, the coders were advised to use a decision rule whereby the presence of more than five criteria equals a good indication of high credibility, and using decision rules is something that CBCA coders should not do. If I disregard their findings, Table 8.3 reveals that accuracy rates for truths varied between 44% to 91% and accuracy rates for lies between 60% and 100%. The average accuracy rate for truths (70.81%) is similar to the accuracy rate for lies (71.12%). The average total accuracy rate in those studies is 70.47%.[14] In studies where direct comparisons were made, accuracy rates for children and adults and for witnesses, victims, and suspects did not appear to differ from each other, further supporting the suggestion that CBCA assessments are not restricted to children's or witnesses' statements.[15]

VALIDITY CHECKLIST: RESEARCH FINDINGS

To date, Validity Checklist research has concentrated on the impact of three issues on CBCA scores: age of the interviewee; interviewer's style; and coaching of the interviewee. I will review these studies in this section.

[14] Before calculating these accuracy rates, I averaged the three scores of Akehurst *et al.* (2001) as well as the two scores of Ruby and Brigham (1998), Vrij, Akehurst *et al.* (2004a), and Vrij, Kneller and Mann (2000). Therefore, all studies had equal weight in the accuracy calculations.

[15] To my knowledge, Ruby and Brigham (1998) are the only researchers to date to have examined the impact of participants' ethnicity on CBCA scores. Although they found no differences in accuracy rates, the ethnicity issue merits attention in future studies given potential differences in narrative techniques between people belonging to different cultures (Davies, 1994b; Phillips, 1993; Vrij & Winkel, 1991, 1994).

Age of the Interviewee

As I mentioned above, cognitive ability, command of language, and memory retrieval strategies develop throughout childhood, making it gradually easier to give detailed accounts of what has been witnessed (Davies, 1991, 1994a; Fivush, 2002; Fivush *et al.*, 1995; Saywitz, 2002). Therefore, the statements of younger children are likely to contain fewer details than statements of older children and adults. Also, children under eight years old may have difficulty in viewing the world from somebody else's perspective (Flavell, Botkin, Fry, Wright, & Jarvis, 1968), thus Criterion 13 (*attribution of perpetrator's mental state*) is unlikely to occur in the statements of young children. Finally, younger children have less developed meta-cognitive and meta-memorial capabilities (i.e., knowing whether or not they know or remember an answer, Ghetti & Alexander, 2004; Magnussen *et al.*, 2006; Walker & Warren, 1995), so are less likely to be aware of gaps in their memories (Criterion 15). Research has convincingly demonstrated that total CBCA scores are age dependent: with increased age, children obtain higher total CBCA scores.[16]

In three field studies the frequency of occurrence of each individual CBCA criterion was examined as a function of age (Boychuk, 1991; Buck, Warren, Betman, & Brigham, 2002; Lamers-Winkelman & Buffing, 1996). Those studies found that many of the 19 criteria are age-related. For example, Buck *et al.* (2002) examined sexual abuse interviews with children aged 2 to 14 years and found that 13 of the 19 criteria were correlated with age. Each of the six criteria that were not correlated (*unusual details*, *accurately reported details misunderstood*, *attribution of perpetrator's mental state*, *raising doubts about one's own memory*, *self-deprecation*, and *pardoning the perpetrator*) were present in less than 10% of the interviews.

More field researchers have examined the frequency of occurrence of criteria in statements regardless of age. They found that the frequency differs widely per criterion.[17] In particular, Criterion 1 (*logical*

---

[16] Anson, Golding, & Gully, 1993; Blandon-Gitlin, Pezdek, Rogers, & Brodie, 2005; Boychuk, 1991; Buck, Warren, Betman, & Brigham, 2002; Craig *et al.*, 1999; Davies, Westcott, & Horan, 2000; Hershkowitz, Lamb, Sternberg, & Esplin, 1997; Horowitz, Lamb, Esplin, Boychuk, Krispin, & Reiter-Lavery, 1997; Lamb *et al.*, 1996; Lamers-Winkelman & Buffing, 1996; Pezdek *et al.*, 2004; Santtila *et al.*, 2000; Vrij, Akehurst *et al.*, 2002, 2004a, b. Some studies did not obtain significant age effects (Akehurst *et al.*, 2001; Tye *et al.*, 1999). However, in Tye *et al.*'s (1999) study, children's ages were not balanced for true and false statements. The correlation between age and total CBCA score in Hershkowitz *et al.*'s (1997) study was only marginally significant ($p < .10$).

[17] Anson *et al.*, 1993; Boychuk, 1991; Buck *et al.*, 2002; Esplin *et al.*, 1988; Horowitz, Lamb *et al.*, 1997; Lamb *et al.*, 1997b; Lamers-Winkelman & Buffing, 1996. See Vrij, 2005b, for an overview of the frequency percentages obtained in these field studies.

structure), Criterion 3 (*quantity of details*), Criterion 4 (*contextual embeddings*), and Criterion 19 (*details characteristic of the offence*) are often present, whereas Criterion 10 (*accurately reported details misunderstood*), Criterion 16 (*raising doubts about own memory*), and Criterion 17 (*self-deprecation*) rarely occur in statements (typically in less than 10% of the statements). This may well explain why the latter three criteria show the least support for the Undeutsch hypothesis (Appendix 8.1). By definition, a criterion that is never present in a statement cannot discriminate between true and false accounts; and a criterion that is hardly ever present is unlikely to discriminate between true and false statements.

Interview Style

CBCA scores are also related to interview style.[18] For example, open-ended questions ("Could you please tell me what happened?") and facilitators (non-suggestive words of encouragement, Hershkowitz, Lamb, Sternberg, & Esplin, 1997) yield more CBCA criteria than other more direct forms of questioning, such as focusing the attention of the interviewee on details that he or she has previously mentioned (Craig *et al.*, 1999; Hershkowitz *et al.*, 1997). Positive correlations between CBCA scores and verbal affirmations ("Yes, I see", etc.) and confirming comments (i.e., interviewer summarising what the child has said) were also found (Davies, Westcott, & Horan, 2000).

Special interview techniques are designed to increase the amount of information that can be obtained from an interviewee. One of those techniques is called the *Cognitive Interview* and was developed by the American psychologists Ron Fisher and Ed Geiselman (1992). Nowadays this technique is often used by police officers in Germany, England, and the United States (and perhaps in other countries) for interviewing cooperative witnesses. People are sometimes unable to spontaneously retrieve information that is stored in their memory. The Cognitive Interview is based upon psychological principles of remembering and retrieving information from memory, and cooperative witnesses who are interviewed with this technique provide more accurate information than those who are interviewed with conventional interview techniques (Köhnken, Milne, Memon, & Bull, 1999; Milne & Bull, 2003). Furthermore, witness statements obtained with the Cognitive Interview receive

[18] Craig *et al.*, 1999; Davies, Westcott, & Horan, 2000; Hershkowitz, 1999; 2001b; Hershkowitz, Lamb, Sternberg, & Esplin, 1997; Köhnken *et al.*, 1995; Lamb, Esplin, & Sternberg, 1995; Lamb, Hershkowitz, Sternberg, Esplin, Hovav, Manor, & Yudilevitch, 1996; Santtila *et al.*, 2000; Steller & Wellershaus, 1996; Sternberg, Lamb, Hershkowitz, Esplin, Redlich, & Sunshine, 1996; Vrij, Mann, Kristen, & Fisher, 2007.

higher CBCA scores than statements obtained with conventional interview techniques (Köhnken *et al.*, 1995; Steller & Wellershaus, 1996). See Fisher, Brennan, and McCauley (2002) for more on the Cognitive Interview.

Coaching the Interviewee

If an interviewee realises that CBCA will be used to assess their credibility, they may attempt to incorporate CBCA criteria in their statements so that they will make a credible impression on coders. Such attempts are called *countermeasures*. Although in principle both truth tellers and liars could attempt to employ countermeasures, liars are most likely to do so, as they are less inclined than truth tellers to take their own credibility for granted (DePaulo, Lindsay *et al.*, 2003).

I am aware of five CBCA countermeasures experiments (Caso, Vrij, Mann, & DeLeo, 2006; Joffe & Yuille, 1992; Vrij, Akehurst *et al.*, 2002, 2004b; Vrij, Kneller, & Mann, 2000). In our experiments, participants of different age groups (6–8 year olds, 10–12 year olds, 14–15 year olds, adults) spoke the truth or lied about several events. Prior to this, some of those participants (both truth tellers and liars) were informed about several CBCA criteria. After receiving this information these “coached” participants were instructed to include those CBCA criteria in their truthful or deceptive statements. The statements were transcribed and then evaluated by CBCA coders. The CBCA scores of the coached participants were compared with the CBCA scores of naive participants who did not receive information about CBCA. Coaching made CBCA an inefficient veracity detection tool. Typically, coached participants obtained higher CBCA scores than naive participants and differences in CBCA scores between truth tellers and liars only emerged in naive participants. Hence, it is difficult to correctly identify coached liars with CBCA analyses. For example, in our study with adult participants (Vrij, Kneller, & Mann, 2000), a CBCA expert correctly classified 69% of the naive liars but only 27% of the coached liars. When we informed the expert that some of the liars had been coached, and then asked the expert to rate the statements a second time, the accuracy rates for coached liars increased, but only to 40%. This means that even after the CBCA expert was informed that some liars had been coached, the majority of liars (60%) were still classified as truth tellers by the CBCA expert.

Although coaching enabled participants to lie successfully in our experiments, 6–8 year olds were the exception. They were not able to enhance the quality of their statements after being informed about CBCA, and this may have been caused by the nature of the coaching. We only gave limited information and the children had little time to practise what they had just learned. Also, the coach was a person unknown to the child. I would not rule out the possibility that even young children can be successfully coached if the coach is someone they know and trust and if the coach practises the rehearsed story extensively with the child. Our experiments, however, do show that it is more difficult to coach younger children than older children or adults.

BELIEFS ABOUT THE RELATIONSHIP BETWEEN CBCA CRITERIA AND TRUTH TELLING

The previous section showed that liars who are informed about the CBCA method can incorporate CBCA criteria into their statements in order to fool CBCA experts. Given people's capability to adjust their statements in order to sound credible, it is important to know what people who are unaware of the CBCA method think the relationship is between the CBCA criteria and truth telling. If people who are unaware of the CBCA method are highly motivated to make a convincing impression, they may attempt to produce statements that they believe sound credible. Are these statements likely to be believed by CBCA experts? They would if people's views about how credible stories sound show overlap with how CBCA experts believe such stories sound. Little research has been conducted that addresses this issue. I am only aware of four published surveys (see Table 8.4). In those surveys, participants were provided with a list of CBCA criteria and asked whether they believed that each of these criteria would occur: (i) more often or (ii) less often during fabrication compared to truth telling or (iii) does not differ between fabrication and truth telling. Three studies were conducted in the United Kingdom with adults, and included both professional lie catchers (police officers, teachers, social workers) and laypersons as participants (Akehurst *et al.*, 1996; Taylor & Vrij, 2000; Vrij, Akehurst, & Knight, 2006). The findings for these groups were similar, indicating that professionals and laypersons have similar beliefs about verbal characteristics of truth telling. The results in Table 8.4 represent the views of all these groups combined.

In Table 8.4 a ">" sign indicates that participants believed that the criterion occurs more often in true accounts than in false accounts, whereas a "<" sign indicates that they believed that the criterion occurs less often in true than in false accounts. A "–" sign indicates that participants did not consider that the criterion is related to truth telling or lying, whereas a "*" indicates that the relationship between the

Table 8.4 People's beliefs about Criteria–Based Content Analysis Criteria and truth telling

	CBCA criteria																		
Authors	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Akehurst, Köhnken, Vrij, & Bull (1996)	*	*	>	<	>	–	–	<	<	<	<	*	*	<	>	<	<	<	–
Colwell, Miller, Miller, & Lyons (2006)	>	<	>	*	*	*	*	*	*	*	*	*	*	<	<	*	*	*	*
Taylor & Vrij (2000)	–	–	–	<	*	*	<	<	<	*	*	*	*	<	<	>	>	*	*
Vrij, Akehurst, & Knight (2006)[1]	>	<	>	*	>	–	–	<	<	*	*	>	*	–	<	*	*	*	*

> observers believe that the CBCA criterion occurs more often during truth telling than when lying; < observers believe that the CBCA criterion occurs less often during truth telling than when lying; – observers do not associate the CBCA criterion with truth telling/lying; * verbal characteristic was not investigated
The Undeutsch hypothesis predicts that all criteria occur more frequently during truth telling than when lying

Note: [1]Adults' beliefs only

criterion and truth telling/lying was not investigated in that particular study. As mentioned above, CBCA experts assume that all criteria occur more frequently during truth telling than fabrication (a ">" sign in the table), and CBCA studies often reveal this pattern.

Table 8.4 shows that people believe that truth tellers' stories contain more *details* (Criterion 3) and more *descriptions of interactions* (Criterion 5) than stories of liars. This is in alignment with CBCA research. For many other cues, people do not know how they actually relate to truth telling. In fact, they often express beliefs that are the direct opposite of reality. For example, people believe that liars include more *contextual embeddings* (Criterion 4), *unusual details* (Criterion 8), and *superfluous details* (Criterion 9) in their stories, whereas, in fact, truth tellers are more likely to do this.

In summary, people's beliefs about how credible statements sound differ to some extent from the views of CBCA experts. This is good news for these experts, because this makes it less likely that liars who deliberately attempt to produce convincing statements without knowing the CBCA tool will provide statements that appear credible to CBCA experts.

Several other studies have focused on the relationship between perceived credibility and one particular CBCA criterion: *number of details*. Most studies support the finding of the CBCA surveys that inclusion of many details makes an honest impression (Bell & Loftus, 1988, 1989; Conte, Sorenson, Fogarty, & Rosa, 1991; Wells & Leippe, 1981).

However, Freedman, Adam, Davey, & Koegl (1996) found that this relationship is dependent on context. When observers are suspicious that somebody is lying, maximum impact is incurred by imparting an intermediate level of detail. To give too many details in these circumstances could easily give observers the impression of trying too hard to be convincing, being too defensive, or making up some of the material. Also Coolbear (1992) reported that too much detail given by a child evokes suspicion, because it suggests that the child has been coached in the story by others. It thus seems that stories sound suspicious if they violate what are perceived as acceptable norms. Similar findings were obtained when examining the relationship between perceived credibility and nonverbal cues (Chapter 5). That research revealed that intermediate levels of behaviour, such as intermediate levels of eye contact, make a more credible impression than norm-violating behaviours, such as gaze aversion or staring.

VALIDITY CHECKLIST: REFLECTIONS

Research into the use of the Validity Checklist in real-life cases is rare, and only three studies which did so have been published (Gumpert & Lindblad, 1999; Lamers-Winkelman, 1999; Parker & Brown, 2000). I found the study conducted by Gumpert and Lindblad particularly interesting and will discuss their work below. First, I will raise some concerns about the Validity Checklist and base those on psychological principles and research.

Difficulty in Identifying Validity Checklist Issues

Some Validity Checklist issues might be difficult to identify. For example, SVA experts look for evidence that an adult has coached a child to enhance the perceived credibility of his or her statement. However, I already discussed that coaching may be difficult to detect for CBCA experts.

Difficulty in Measuring Validity Checklist Issues

At least one Validity Checklist issue, *susceptibility to suggestion* (Criterion 3, Steller, 1989), is difficult to measure. Some witnesses are more prone to an interviewer's suggestions than others. The danger of suggestibility is that a suggestible child may be inclined to provide information that confirms the interviewer's expectations but that, in fact, is inaccurate. Yuille (1988b) and Landry and Brigham (1992) therefore recommend asking the witness a few misleading questions at the end of the interview to assess the witness' susceptibility to suggestion. Since asking such questions about central information could harm the statement (I discussed this above), they recommend focusing on peripheral information (e.g., "When you were with your sister, which friend was also there, Claire or Sarah?" when the interviewer knows that there was no friend present). Being restricted to asking questions about peripheral information is problematic. Children show more resistance to suggestibility for central than for peripheral parts of an event (Dalton & Daneman, 2006; Goodman, Rudy, Bottoms, & Aman, 1990), and are more resistant to suggestibility for stressful events, most likely the central events, than for less stressful events, most likely the peripheral events (Davies, 1991). Therefore, insight into children's suggestibility for peripheral parts of the event cannot be effectively used to draw conclusions about their suggestibility for core events. This Validity Checklist issue also seems to be based on the assumption that suggestion is more the result of individual differences than of circumstances. This may not be a valid assumption (Milne & Bull, 1999).

Difficulty in Determining the Impact of Validity Checklist Issues

Once Validity Checklist issues are identified and measured it is often difficult, if not impossible, to determine their exact impact on CBCA scores. I already mentioned that interviewing style often influences the amount of information that witnesses report. However, this does not mean that each individual witness will necessarily be affected by a particular interview style. And even if the witness was affected by the style of interviewing, it is often impossible to determine afterwards the precise impact this had on the quality of the statement. In other words, the precise impact of Validity Checklist issues on an individual statement can often not be determined, but only *estimated*. Obviously, mistakes in these estimations may occur.

A good illustration of the difficulty that SVA experts face in determining the exact impact of Validity Checklist issues on CBCA scores is the field study conducted by Lamers-Winkelman and Buffing (1996). In this study, raters were instructed to take the age of the child into account when calculating CBCA scores. Nevertheless, six criteria positively correlated with age. In other words, even after being instructed to correct CBCA scores for age, the results still showed age-related effects with older children obtaining higher CBCA scores than younger children.

Given these difficulties in identifying and measuring the presence of Validity Checklist issues, and in examining the exact impact of these issues on CBCA scores, it is clear that the Validity Checklist procedure is more subjective and less formalised than the CBCA procedure (Steller, 1989; Steller & Köhnken, 1989). It is therefore not surprising that if two experts disagree about the veracity of a statement in a German criminal case, the most likely reason for the disagreement is that they disagree about the impact of Validity Checklist issues on that statement (Köhnken, 1997, personal communication). Also, when examining the use of the Validity Checklist in Sweden, Gumpert and Lindblad (1999) found that different experts sometimes draw different conclusions about the impact of Validity Checklist issues on children's statements. It is therefore preferable that in criminal investigations not one but at least two evaluators assess a case independently from each other. Presently, it is not common practice to ask for a second opinion.[19]

Justification of Validity Checklist Issues

I question the justification of some of the Validity Checklist issues, particularly Issue 2 *inappropriateness of affect* (Steller, 1989), Issue 9 *consistency with the law of nature* (Steller, 1989), Issue 10 *inconsistency with other statements* (Steller, 1989), and Issue 11 *consistency with other evidence* (Steller, 1989).

Inappropriateness of affect

Issue 2 refers to whether the child displayed inappropriate affect during the interview (Raskin & Esplin, 1991b). It suggests that if a child reports details of abuse without showing any signs of, or inappropriate signs of, emotion then the story might be less trustworthy. This is a view that is also commonly expressed by non-SVA experts. Coolbear (1992) found that professionals from legal professions and human services thought that emotional congruence with the nature of the material covered by the child indicates truthfulness. I already mentioned

[19]The problem that SVA experts have to deal with, a witness' response is not just influenced by the veracity of a statement but also by other issues, is not unique to SVA assessments. Lie detectors who examine behavioural or physiological responses also face this problem. They attempt to resolve this by introducing a baseline response: a typical, natural response of the interviewee which the lie detector knows to be truthful and which is provided in similar circumstances to the response under investigation. They then compare the baseline response with the response under investigation, and, since they created a setting in which the impact of external factors on both responses is assumed to be the same, differences between the two responses might indicate deception. However, the method is complex, as creating a good baseline is often problematic. I have already addressed this baseline issue in Chapter 3 and will return to it later in this book.

in Chapter 6 that the 14-year-old Michael Crowe became a prime suspect because the detectives believed that he had reacted to his sister's death with "inappropriately little emotion" (Kassin, 2005); and I reported in Chapter 5 that German police detectives, who specialised in the investigation of sexual assault, become suspicious if the victim displays "atypical behaviour".

My concern is that this incorrectly assumes that atypical behaviour/inappropriate affect or, the opposite, typical behaviour/appropriate affect exists. Research with rape victims has distinguished two basic styles of self-presentation: An "expressed" style in which the victim displays distress which is clearly visible to outsiders, but also a more controlled "numbed" style, whereby cues of distress are not clearly visible (Burgess, 1985; Burgess & Homstrom, 1974; Vrij & Fischer, 1995). The varying communication styles represent a personality factor, and are not related to deceit (Littman & Szewczyk, 1983). Yet, different emotional displays have a differential impact on the perceived credibility of victims, and emotional victims are more readily believed than victims who report their experience in a controlled manner (Baldry & Winkel, 1998; Baldry, Winkel, & Enthoven, 1997; Bothwell & Jalil, 1992; Kaufmann, Drevland, Wessel, Overskeid, & Magnussen, 2003; Vrij & Fischer, 1997; Wessel, Drevland, Eilertsen, & Magnussen, 2006; Winkel & Koppelaar, 1991). Given that inappropriate affect does not exist and that people tend to draw sometimes incorrect conclusions on the basis of the displayed affect, it is unfortunate to encourage SVA evaluators to pay attention to it. For this matter, I like more the consultation paper from the UK solicitor general Mike O'Brien in which he argues that victims who appear unusually composed in the witness box should no longer be treated with suspicion (*The Observer*, 26 March 2006, p. 19).

Inconsistencies between statements

Issue 10 deals with inconsistencies between different statements from the same witness. It suggests that fabrications may have occurred when interviewees contradicted themselves in two different statements. However, when reviewing child research, Fivush, Peterson, and Schwarzmueller (2002) found that inconsistency between different statements is not a valid indicator of deception.[20] Moreover, this child research has demonstrated that children's narratives naturally change across different interviews, which is largely due to differences in interviewers across interviews and the type of questions asked.

[20] Inconsistency between different statements is not a valid indicator of deceit in adults either, as Chapter 4 revealed.

Köhnken (2004) acknowledges the problems associated with examining inconsistencies and argues that looking at inconsistencies per se is inappropriate. Rather, he believes that looking at "substantial discrepancies . . . in the description of the core of an event . . . would at least require an explanation" (Köhnken, 2004). This cautious conclusion makes sense, but still may be problematic, particularly when evaluating children's statements. Children can easily give conflicting statements even about core events, for several reasons including that of question repetition (Moston, 1987; Poole & White, 1991; Quas, Davis, Goodman, & Myers, 2007). Young children in particular may give two different answers to the same question, because they may think "If the same question is asked on a second occasion, the first answer I gave must have been wrong." Their reasoning is understandable because this often happens to them. Parents and teachers often repeat a question when the initial answer the child gave was wrong. Since children are often repeatedly interviewed in sexual abuse cases, and are often asked the same questions in different interviews, problems thus may arise when examining inconsistencies.

Realism of the statement

Issues 9 and 11 deal with the question of whether the event described in the statement is unrealistic or impossible. If this is the case, it should raise doubts about the truthfulness of the statement according to SVA experts. However, the presence of unrealistic or impossible details in a statement does not necessary indicate that the statement was fabricated. Dalenberg, Hyland, and Cuevas (2002) examined the statements of children who made initial allegations of abuse and for whom there was a "gold standard" of proof that abuse had occurred (e.g., the injuries were judged medically consistent with the allegations). They found that for a small group of children bizarre and improbable material was included in their statements such as reference to fantasy figures, impossible or extremely implausible features of the story, and descriptions of extreme abusive acts that should have been (but were not) supported by external evidence if they had genuinely occurred. Since those statements contain unrealistic elements, there is a risk that they would be considered as untrue on the basis of a Validity Checklist assessment.

Some Relevant Validity Checklist Issues are not in the Validity Checklist

Not all issues that are known to influence CBCA scores are present in the Validity Checklist. As a result, their influence will be ignored by SVA experts. For example, when interviewees talk about events they are familiar with, their statements include more CBCA criteria then when they talk about unfamiliar events (Pezdek *et al.*, 2004). The interviewee's familiarity with the event may even have a stronger impact on CBCA scores than the actual veracity of the statement (Blandon-Gitlin, Pezdek, Rogers, & Brodie, 2005). The reason why talking about familiar events results in high CBCA scores, even when the statement is false, is that witnesses can draw upon their experience with such events when making up a false statement. Raskin and Esplin (1991b) acknowledge this problem. They report that CBCA is difficult to apply in "situations where the witness has other sources of information from which to invent an accusation that incorporates some or many of the content criteria" (Raskin & Esplin, 1991b, p. 280). Pezdek suggested that in real sexual abuse cases, SVA may be an effective tool in distinguishing between true and fabricated accounts only for people without prior knowledge or experience with sexual acts (Pezdek & Taylor, 2000). The interviewee's familiarity with the event is not examined in the SVA procedure.

Another issue that is important to examine but which is not included in the Validity Checklist is the number of times that a person has been interviewed. Once an alleged victim of child abuse enters the legal system, he or she is typically interviewed in detail on several occasions (Granhag, Strömwall, & Landström, 2006). Research has shown that the number of times someone is interviewed affects the quality of a statement. Children may provide more details in a second interview than in the first interview (Goodman & Schwartz-Kenney, 1992; Yuille & Cutshall, 1989). Perhaps children feel uncomfortable during the first interview with an interviewer they do not know and therefore dare not to say much, or perhaps they need to build up trust with the interviewer and therefore do not wish to say much in the first interview. However, Boychuk's (1991) field study showed that conducting too many interviews has a negative impact on the information provided by interviewees. She compared the CBCA scores of children who were interviewed for the first, second, third, fourth, or more time. She found that the statements of children who had already been interviewed at least three times yielded lower CBCA scores than the statements of the other children. Boychuk suggested that after a couple of interviews a child may become tired of talking about the topic.

In their laboratory experiment, Erdmann, Volbert, and Böhm (2004) examined whether the number of interviews conducted with a child affects the accuracy of CBCA assessments. They interviewed 6–8-year old participants five times about real and fictitious events. CBCA analyses were carried out after the first and fifth interview. Although CBCA scores did differ between true and fictitious accounts after the first interview, very few differences emerged after the fifth interview, suggesting that it becomes increasingly difficult to discriminate between true and fictitious accounts via the CBCA method as the number of times that someone is interviewed increases. Taken together, these findings suggest that CBCA analyses are most effective if they are carried out on statements provided during the initial interviews.

Research has further shown that CBCA scores are related to verbal and social skills (Santtila *et al.*, 2000; Vrij, Edward, & Bull, 2001c; Vrij, Akehurst *et al.*, 2002, 2004b). For example, Santtila *et al.* (2000) found a positive correlation between CBCA scores and verbal ability assessed with the WISC-R Vocabulary test; and we found that in some age groups CBCA scores were positively correlated with social adroitness and self-monitoring, and negatively correlated with social anxiety (Vrij, Akehurst *et al.*, 2002). Measuring verbal and social skills is not part of the Validity Checklist.

Other issues, presently unknown, may also impact on CBCA scores. For example, some children have psychological disorders, such as depression or attention problems, and it is unclear what impact these may have on their CBCA scores.[21]

The Validity Checklist Might be Improperly Used

Gumpert and Lindblad's (1999) field study regarding the use of the Validity Checklist by Swedish SVA experts revealed that the experts sometimes use the Validity Checklist incorrectly, possibly due to difficulties in applying it. First, although SVA experts sometimes highlight the influence of Validity Checklist issues on children's statements in general, they do not always discuss how these issues may influence the statement of the particular child they are asked to assess. Second, although experts sometimes indicate possible external influences on statements, they are inclined to rely upon the CBCA outcome, and tend to judge high quality statements as truthful and low quality statements as fabricated.

Gumpert and Lindblad (1999) only examined a limited number of cases, and drawing conclusions is therefore perhaps premature. Their findings, however, are worrying if they could be replicated with a larger sample. First, they imply that Validity Checklist issues tend to be

[21] The difficulty in measuring Validity Checklist issues and determining their exact impact on CBCA scores, discussed above, also applies to these concepts. For example, how should social adroitness and self-monitoring be assessed in the individual case? (In this respect, research demonstrating that people's natural language use is an indicator of their personality is relevant, Pennebaker & Graybeal, 2001; Pennebaker & King, 1999). And how can the extent to which social adroitness and self-monitoring influenced the quality of the statement be determined in an individual case?

ignored. It is easy to explain why this happens. As already discussed, it is often impossible to judge the precise impact of a Validity Checklist issue on the quality of a statement, and just disregarding its impact is often the easiest solution. Second, ignoring the influence of Validity Checklist issues means that SVA decisions are not likely to be more accurate than CBCA assessments, as the final decision based upon CBCA outcomes together with the Validity Checklist procedure will often be the same as the decision based upon CBCA outcomes alone. Third, ignoring the influence of Validity Checklist issues implies that interviewees who will naturally produce low quality statements and therefore most likely obtain low CBCA scores (e.g., young children, interviewees with poor verbal skills) may be in a disadvantageous position if their extenuating circumstances are ignored.

Difficulty in Detecting Embedded False Statements/Lies and False Memories

A child may have been sexually abused and accurately report his or her experiences, but may inaccurately (deliberately or not) accuse an innocent person as the perpetrator. This false allegation runs the risk of being judged as truthful. Since the inaccuracy is embedded in an accurate statement, most elements of this false statement, which I would like to call an *embedded false statement* if the wrong person was accused by mistake and an *embedded lie* if accused on purpose, are true. Hence, such a statement could well be rich in detail and achieve a high CBCA score.

Moreover, sometimes people, both adults and children, are confused about what they have actually experienced and what they have only imagined (Foley & Johnson, 1985; Johnson & Foley, 1984; Markham, 1991; Parker, 1995). Their narratives of these imagined events may be detailed, and therefore likely to obtain high CBCA scores, and possibly be judged as truthful. Several studies have shown how detailed memories can be of imagined, non-experienced, events. Starting with adults, a compelling example was given by Crombag, Wagenaar, & Van Koppen (1996). Their study was based on a real-life event, the crash of a cargo El Al Boeing 747 into an 11-storey apartment building in Amsterdam (The Netherlands) on 4 October 1992. Dutch television reported extensively on this national disaster, showing the fire brigade fighting the blaze and rescuing people from the collapsing building. The disaster was the main news item for several days and eventually everyone in the country knew – or thought they knew – in great detail what had happened. In all these news programmes, the television could only show what had happened after the crash; the crash itself was never filmed and thus never broadcast. Nevertheless, when asked "Did you see the television footage of the moment the plane hit the apartment building?" 61 (66%) of the 93 undergraduate students who participated in the study responded that they had. Many "witnesses" provided further details of this non-existent television footage of the crashing plane. For instance, 41 participants remembered that they had seen the plane hit the building horizontally, 10 remembered that the plane hit the building vertically, and 14 indicated that they had seen on television that the plane was already in flames when it crashed. In summary, many participants had rich and vivid memories about an event they had never witnessed, but which they thought they had witnessed.

In a similar vein, we asked adult participants whether they had seen the non-existent film of the car crash in Paris in which Princess Diana, Dodi Fayed, and their driver were killed (Ost Vrij, Costall, & Bull, 2002). Of the 45 participants who were asked this question, 20 (45%) said they had. To give a final striking adult example, Johnson, Hashtroudi, and Lindsay (1993) described a CBS television programme in which the then US President Reagan recounted a story to Navy personnel about an act of heroism that he attributed to a real US pilot. However, no record of this real act or a similar act of heroism could be found, but the story bore an uncanny resemblance to a scene from a Dana Andrews movie released in the 1940s.

In their experiment with children, Ceci and his colleagues asked 3 to 6 year olds to imagine several events, including events that had never happened to them, such as falling off a tricycle and getting stitches in their leg (negative event) and going on a hot air balloon ride with their classmates (positive event) (Ceci, Loftus, Leichtman, & Bruck, 1994). They held weekly interviews with the children about the events they had imagined. During the eleventh interview, 59% of the three and four year olds and 51% of the five and six year olds thought that they actually remembered having once been in a hot air balloon and 31% of the three and four year olds and 28% of the five and six year olds told the interviewer that they had actually fallen off their tricycle and received stitches in the leg. All of the interviews were videotaped, and these tapes revealed internally coherent and detailed narratives. They are therefore likely to obtain high CBCA scores and to be judged as truthful. Yet, they were false. In follow-up studies, Ceci and his colleagues showed videotapes of the children's eleventh session to professionals (clinicians and researchers who were specialised in interviewing children, but not trained in CBCA) to see if they could determine which events had actually been experienced by the children and which were fictitious (Ceci, Huffman, Smith, & Loftus, 1994; Ceci, Loftus *et al.*, 1994). These professionals were no better than chance at distinguishing between accurate and inaccurate reports. It is unfortunate that the observers in the Ceci studies were not trained in using CBCA, because it would be interesting to examine whether CBCA experts would fare better than non-CBCA trained professionals.

The closest to examine whether CBCA can discriminate between real and imagined events was Steve Porter and his colleagues (Porter, Yuille, & Lehman, 1999). Their study with adult participants is already described in Chapter 2 (Box 2.1). Via guided interviewing they "implanted" memories of events that never occurred in their (adult) participants (e.g., being sent to a hospital emergency room). As a result, many participants believed they had experienced these events whereas they, in fact, had never experienced them. The researchers compared the accounts of these implanted events with the accounts of events that the participants had truly experienced. They examined, amongst other cues, two CBCA criteria, *quantity of details* (Criterion 3) and *admitting lack of memory* (Criterion 15). No differences in the frequency of occurrence of these two criteria were found between the implanted and truly experienced accounts. Although one study investigating only two CBCA criteria is not enough to draw any conclusions, it sounds plausible that CBCA coders will have difficulty in correctly classifying false accounts that are the result of false memories. Not only CBCA experts will have difficulty in determining the veracity of such accounts, I already discussed in Chapter 3 that false beliefs are also unlikely to reveal behavioural cues to deceit, simply because people are not lying in such instances. However, there is some evidence that false memories can be detected via Reality Monitoring, as I will discuss in Chapter 9.

Box 8.3 Strategy used by non-SVA users

Coolbear (1992) investigated to what extent professionals who are *not* familiar with SVA actually use criteria similar to the Validity Checklist in assessing the credibility of allegations. She held structured interviews with 51 professionals from legal professions and human services experienced in child sexual abuse allegations. The most common response was that the use of childlike language was an indicator of truthfulness (Issue 1 on the Validity Checklist). Emotional congruence with the nature of the material discussed by the child (Issue 2 of the Validity Checklist) was also mentioned as indicative of a truthful story. Finally, many participants stated that they would proceed cautiously with a sexual abuse allegation if they knew that there was a custody/access dispute or divorce process between the child's parents (Issues 6 to 8 on the Validity Checklist). In other words, the strategy used by these professionals resembled to some extent the approach used by SVA experts.

LEGAL IMPLICATIONS

In this section I discuss what are, in my view, the implications of this chapter's findings for the use of CBCA/SVA assessments as scientific evidence in criminal courts. Throughout this book, I will include these sections for veracity assessment tools that are used as evidence in criminal courts or that should be made admissible in criminal courts according to their supporters. I address this issue of admissibility by examining whether the tools meet the criteria that are required for admitting expert scientific evidence in criminal courts. I will hereby use the set of guidelines provided by the United States Supreme Court for admitting expert scientific evidence in (American) federal courts. These guidelines were presented in the *Daubert v. Merrel Dow Pharmaceuticals, Inc.* (1993) case. The following five criteria were used in the *Daubert* case (Honts, 1994): (a) Is the scientific hypothesis testable? (b) Has the proposition been tested? (c) Is there a known error rate? (d) Has the hypothesis and/or technique been subjected to peer review and publication? and (e) Is the theory upon which the hypothesis and/or technique is based generally accepted in the appropriate scientific community? Table 8.5 summarises my answers to these questions for SVA assessments.

Question 1: Is the Scientific Hypothesis Testable?

The prediction that truthful statements will obtain higher CBCA scores than false statements (the Undeutsch hypothesis) can easily be tested in laboratory research. Although testing the Undeutsch hypothesis in field studies is possible in principle, in reality it is difficult given the problems with establishing the truth or falsity of statements beyond doubt (ground truth). My answer to the first *Daubert* question is therefore "yes" for CBCA laboratory research but "problematic" for CBCA field research. Some underlying assumptions of the Validity Checklist are also difficult to test in real life. For example, how can the extent to which interview style has influenced the quality of the statement be established? The answers are therefore "problematic" for the Validity Checklist and for SVA as a whole.

Table 8.5 Answers to the five *Daubert* questions for CBCA and SVA assessments

	CBCA laboratory	CBCA field	Validity Checklist	SVA
(1) Is the scientific hypothesis testable?	yes	problematic	problematic	problematic
(2) Has the proposition been tested?	yes	no	no	no
(3) Is there a known error rate?	yes, too high	no	no	no
(4) Has the hypothesis and/or technique been subjected to peer review and publication?	yes	yes	no	no
(5) Is the theory upon which the hypothesis and/or technique is based generally accepted in the appropriate scientific community?	unknown	unknown	unknown	unknown

Question 2: Has the Proposition Been Tested?

My answer to the second *Daubert* question is affirmative for CBCA laboratory research. Appendix 8.1 lists 32 laboratory samples in which the Undeutsch hypothesis has been tested, although in most studies adults rather than children participated. There are fewer field studies conducted and several of them are of poor quality. I am therefore inclined to answer the second question with "no" for CBCA field research. Research regarding the Validity Checklist is virtually non-existent, hence the answers for the Validity Checklist and SVA research as a whole are therefore also "no".

Question 3: Is there a Known Error Rate?

There is a known error rate for CBCA judgements made in laboratory research. This error rate is almost 30% for both truths and lies (Table 8.3). Although the error rates indicate that truths and lies can be detected in laboratory studies above the level of chance by using the CBCA tool, it also shows that incorrect judgements are frequently made.

An error rate of 30% implies that CBCA assessments are not made "beyond reasonable doubt", which is the standard of proof typically set in criminal courts. My answer to the third question for CBCA laboratory research is thus "yes, too high". Of particular interest is the error rate of SVA judgements in field studies. A properly conducted field study examining this error rate has not been published to date, so my answers for CBCA field research, the Validity Checklist, and SVA as a whole are all "no".

As long as the error rate in field studies is unknown, there is no better alternative than to use the known error rate in CBCA laboratory studies. This error rate, around 30%, is probably not an unreasonable estimate for the accuracy of SVA judgements. I discussed in the Validity Checklist section that SVA experts in real life tend to rely heavily on the CBCA outcomes when making their final verdicts. Moreover, there is reason to believe that truth/lie assessments in real-life situations are as difficult as, or even more difficult than, truth/lie assessments in laboratory studies, because numerous issues that are typically controlled for in laboratory studies can impact upon CBCA scores in real life. I discussed the difficulty in identifying, measuring, and determining the exact impact of these issues in the Validity Checklist section. In that section I also questioned the justification of some issues that appear in the Validity Checklist and argued that some relevant issues are not included in the Validity Checklist. I also argued that some types of false statements that may appear in real life (embedded false statements/lies and false memories) will be difficult to detect. In summary, although the error rate for SVA assessments in real-life cases is unknown, incorrect decisions are likely to occur given the numerous difficulties associated with making SVA assessments. If I take the known CBCA error rate of 30% as a guideline, it is clear that the accuracy of SVA assessments does not achieve a level that is "beyond reasonable doubt". In other words, SVA assessments are not accurate enough to be presented as scientific evidence in criminal courts.

Question 4: Has the Hypothesis and/or Technique been Subjected to Peer Review and Publication?

A growing number of CBCA studies have now been published in peer-reviewed journals, both laboratory studies and field studies. Validity Checklist studies and SVA studies are still lacking. My answer to the fourth *Daubert* question is thus "yes" regarding CBCA laboratory and field research, but "no" regarding Validity Checklist and SVA research as a whole.

Question 5: Is the Theory upon which the Hypothesis and/or Technique is Based Generally Accepted in the Appropriate Scientific Community?

As already mentioned above, several authors have expressed serious doubts about the CBCA/SVA method (Brigham, 1999; Davies, 2001; Lamb *et al.*, 1997b; Pezdek & Taylor, 2000; Rassin, 1999; Ruby & Brigham, 1997; Wells & Loftus, 1991). For example, Lamb *et al.* (1997b, p. 262) who in my view carried out the only reliable field study to date concluded that "...the level of precision clearly remains too poor to permit the designation of CBCA as a reliable and valid test suitable for the courtroom". However, to answer the fifth *Daubert* question we need to know the extent to which these scholars represent the appropriate scientific community. The appropriate scientific community as a whole has not been consulted about this matter to date. My answer to the final *Daubert* question is thus "unknown".

Overall Verdict

SVA evaluations do not meet the *Daubert* guidelines for admitting expert scientific evidence in criminal courts. In this chapter I have raised doubts about whether the scientific hypothesis underlying the method is actually testable in real life, given the difficulties in establishing the ground truth in field studies. Moreover, not enough field research has been carried out, particularly with regard to the Validity Checklist, and the known error rate in the available CBCA studies makes clear that CBCA assessments are not made beyond reasonable doubt. Finally, it is unknown to what extent the method is accepted in the appropriate scientific community. Regarding the error rate, SVA evaluators may challenge my observation that the error rate is around 30%, as this is the known error rate for *laboratory CBCA* research rather than the error rate for *SVA* evaluations made in *real life*. However, those SVA evaluators must accept that even if those CBCA error rates should be disregarded, the error rate in SVA evaluations is completely unknown. Such an outcome does not meet the *Daubert* guideline either.

EVALUATION OF SVA

The rationale behind SVA is that a statement derived from memory of an actual experience differs in content and quality from a statement based on invention or fantasy, as a result of truth tellers and fabricators experiencing differences in cognitive load and motivation to appear credible. Problems associated with SVA include that the test is not standardised and that issues other than veracity affect the CBCA scores. These issues are often difficult to identify and measure, and their exact impact on CBCA scores can only be estimated. Moreover, some of these issues that SVA experts examine are difficult to justify, whereas other issues that are known to influence CBCA scores are neglected by such experts. Finally, some types of false statements (embedded false statements/lies and false memories) are difficult to detect with the SVA tool.

The accuracy of CBCA (and SVA as a whole) in real-life investigations is unknown since the ground truth was not established in so many field studies, but laboratory studies show that CBCA assessments can detect both truths and lies with around 70% accuracy. The error rate of both truths and lies is thus 30%, which is not beyond reasonable doubt, the standard of proof regularly set in criminal courts. This error rate, which amounts to almost a one in three chance of being wrong, is one of the reasons why in my view SVA assessments do not meet the *Daubert* guidelines for admitting expert scientific evidence in criminal courts, and my main reason for believing that SVA assessments should not be allowed as evidence in criminal courts.

I am aware that this firm standpoint will be challenged by those who propose to allow SVA analyses as evidence in criminal courts. Their main argument is that SVA assessments need to be compared with other evidence presented in criminal courts that may be even less accurate (Köhnken, 2004). I expect that more people working in the criminal justice system will be sensitive to this argument. However, there are arguments against this point of view. For example, I think that people, including the defendant, will perceive the evidence that SVA experts present in court as strong. In cases where the defendant is innocent but the SVA expert considers the accusation made against them is truthful, the defendant may well believe that they can no longer avoid conviction. A false confession to obtain a lower sentence may be the result. In my view, if the evidence is likely to be influential then a high accuracy level is necessary.

However, despite my wishes, if it is decided to allow SVA assessments as evidence in criminal courts then, at the very least, SVA experts should present the problems and limitations associated with their assessments when giving their expert testimony, so that judges, jurors, prosecutors, and solicitors can make a considered decision about the validity of their expert testimony. In addition, because of the subjective nature of the assessments (i.e., human interpretation rather than fact finding), more than one expert should assess each statement to establish inter-rater reliability between evaluators.

Despite these critical comments, I believe that the CBCA list has provided lie detectors with a wealth of verbal criteria they can use to detect deceit. Numerous studies have demonstrated that several of the CBCA criteria, such as *unstructured production*, *quantity of details*, *contextual embeddings*, and *reproduction of conversations* are often diagnostic cues to deceit. It is also striking how consistent the research findings are. Either truth tellers include those, and other, criteria more often in their statements than liars do, or no difference between truth tellers and liars have been found. Truth tellers included fewer of those criteria in their statements than liars in only a few, exceptional, cases. The most convincing findings were obtained when examining a cluster of cues (i.e., total CBCA score), rather than each cue individually. In the vast majority (80%) of studies where a total CBCA score was calculated, truth tellers achieved a higher CBCA score than liars. There is also evidence that the findings are robust. That is, similar findings were obtained in both children and adults; and in both victims/witnesses and suspects, and CBCA discriminated between truth tellers and liars also in contexts where interviewees discussed topics other than sexual abuse. This pattern of findings is far more consistent than that of research into nonverbal cues to deception (Chapter 3) or in research examining verbal cues to deception that are not part of the CBCA list (Chapter 4).

Research findings further suggest that people are also better lie detectors when they examine CBCA criteria compared to when they rely on global assessments of someone's nonverbal or verbal behaviour (Chapter 6). The average accuracy obtained with CBCA analyses was just above 70%, whereas the average truth and lie accuracy obtained by professionals making global assessments of someone's nonverbal and verbal behaviour was just below 55% (Chapter 6). In fact, the *average* accuracy rate obtained in CBCA research is equal to the *highest* accuracy rates achieved in nonverbal and verbal demeanour studies with professional lie catchers. It is therefore a shame that police manuals tend to focus on nonverbal cues to deception or to verbal cues that are not part of the CBCA list, the latter of which is typically neglected in such manuals.

Because truths and lies can be detected above the level of chance with CBCA in a variety of settings and with different type of interviewees I believe that professional lie catchers could use CBCA/SVA assessments in their investigations.[22] For example, in the initial stage of an investigation in order to form rough indications of the veracity of various interviewees; or when different professional lie catchers' opinions are

[22] See Yuille and Cutshall (1989) for a description of an SVA analysis in a murder investigation.

divided about the truthfulness of a particular interviewee. It is hereby necessary to thoroughly train investigators how to interview examinees, since only statements that are in a free narrative style without inappropriate prompts or suggestions from the interviewer are suitable for CBCA assessments. Police detectives, for example, often use a more leading interview style in their interviews with suspects than the SVA procedure allows, which makes CBCA assessments of suspects' statements often inappropriate.

Furthermore, it is necessary to give investigators proper training in how to conduct CBCA assessments, given the erratic effects obtained in studies where trainees were exposed to less comprehensive training programmes. Violating SVA procedures can be disastrous, as Landry and Brigham's (1992) experiment showed. They subjected their observers to a short CBCA training programme, and let them assess statements which were too short. Moreover, these observers watched videotapes rather than analysing transcripts, and made veracity judgements on the basis of decision rules. In that study liars obtained a higher total CBCA score than truth tellers, the only erratic CBCA total score result published to date.

It may also be possible to use CBCA analyses in situations other than a judicial context. CBCA experts disagree about when exactly their analyses can be applied, but some argue that their use is applicable for any statements that relate to events that induce negative emotions, involve a loss of control, and directly involve the interviewee (Steller, 1989; Steller *et al.*, 1988). Many situations outside the judicial context probably fulfil these criteria. Others suggest that CBCA can be used when liars face cognitive difficulty and are motivated to appear credible (Köhnken, 1996, 1999, 2004). This applies to yet more situations.[23]

Obviously, making CBCA assessments is complicated given the requirements of (i) obtaining a statement via a free narrative, and (ii) using transcripts of those statements for coding purposes. It is also time consuming given the many criteria involved. It is therefore perhaps unsuitable for use in many situations. It could be that a quick assessment, using only some criteria, will suffice to obtain a rough idea about the veracity of statements. Research is necessary to examine whether this is the case. Alternatively, a veracity tool that is less time consuming and easier to apply could be used. I will discuss such a tool, Reality Monitoring, in the next chapter.

[23] CBCA evaluations are typically made when people have discussed their own or other people's activities. However, people do not just lie about activities, they also lie about their opinions, attitudes and feelings (Chapter 2). CBCA analyses may well be inappropriate for detecting these types of lies.

Appendix 8.1 Differences between truth tellers and liars on CBCA criteria

Authors	Age	Event	Status	CBCA criteria 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total
Field studies																							
Boychuk (1991)	4–16	field	victim	>	>	>	>	>	>	>	>	>	–	>	>	–	>	–	–	–	>	–	*
Craig *et al.* (1999)	3–16	field	victim	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	>
Esplin *et al.* (1988)	3–15	field	victim	>	>	>	>	>	>	>	>	>	–	>	>	>	>	>	–	–	>	>	>
Lamb *et al.* (1997b)	4–13	field	victim	–	>	>	>	>	>	–	–	–	–	–	–	–	–	*	*	*	*	*	>
Parker & Brown (2000)	adult	field	victim	–	>	>	–	>	>	–	–	–	–	–	>	>	–	–	–	–	–	*	*
Rassin & van der Sleen (2005)	unknown	field	victim	*	*	*	>	>	*	*	–	*	*	*	–	–	*	–	–	*	*	*	*
Laboratory studies																							
Akehurst *et al.* (2001)	7–11/adult	active	n/a	>	–	>	–	>	>	–	–	–	*	*	>	–	–	*	*	*	*	*	>
Akehurst *et al.* (2004)	7–11	active	n/a	–	–	>	>	>	>	–	–	–	*	*	–	*	–	–	–	*	*	*	*
Blandon-Gitlin *et al.* (2005)	9–12	active	n/a	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	–
Caso *et al.* (2006c)[1]	adult	active	suspect	*	*	>	*	>	–	–	*	*	*	*	–	*	–	>	*	*	*	*	*
Colwell *et al.* (2002)	adult	staged	witness	>	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Erdmann *et al.* (2004)[2]	6–8	memory	victim	>	>	>	>	–	–	–	>	–	–	–	>	–	–	–	–	–	–	–	*
Gödert *et al.* (2005)	adult	active	suspect	–	–	>	–	–	>	–	–	>	–	–	–	–	–	–	–	–	–	*	>
Granhag *et al.* (2006)[3]	12–13	staged	witness	–	<	–	*	–	>	*	<	*	*	*	–	*	–	>	>	*	*	*	–
Höfer *et al.* (1996)	adult	active	n/a	>	–	>	>	–	>	>	–	–	*	*	>	–	–	*	*	*	*	*	>
Köhnken *et al.* (1995)	adult	video	witness	–	>	>	*	–	–	–	–	–	–	*	*	–	–	>	–	*	*	*	*
Landry & Brigham (1992)	adult	memory	victim	<	*	>	>	>	>	–	>	>	*	*	>	<	>	>	>	–	*	*	>
Porter & Yuille (1996)	adult	active	suspect	>	*	>	*	*	–	–	–	–	*	–	–	*	–	>	*	*	*	*	*
Porter *et al.* (1999)[4]	adult	memory	victim	*	*	–	*	*	*	*	*	*	*	*	*	*	*	>	*	*	*	*	*
Ruby & Brigham (1998)	adult	memory	victim	<	>	–	<	>	–	>	>	>	*	<	<	*	>	>	–	<	*	*	<
Santtila *et al.* (2000)	7–14	memory	victim	–	>	>	–	–	>	–	>	–	–	–	–	>	–	*	*	*	*	*	*
Sporer (1997)	adult	memory	victim	>	–	–	>	–	–	–	–	–	–	–	–	–	*	*	*	*	*	*	*

Steller *et al.* (1988)	6–11	memory	victim	>	*	>	>	–	–	>	>	>	>	>	–	–	–	–	*	<	–	*	*
Strömwall, Bengtsson *et al.* (2004)	10–13	active	witness	–	–	–	–	–	–	–	–	–	*	–	–	–	>	–	>	–	–	*	–
Tye *et al.* (1999)	6–10	active	witness	–	>	>	>	–	>	–	–	>	*	*	–	–	–	*	*	*	*	*	>
Vrij, Akehurst *et al.* (2002)	5–15/adult	active	witn/sus	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	>
Vrij, Akehurst *et al.* (2004a)	5–15/adult	active	witn/sus	>	*	>	>	>	>	*	*	*	*	*	–	*	–	–	–	*	*	*	>
Vrij, Akehurst *et al.* (2004b)	6–15/adult	active	witn/sus	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	>
Vrij, Edward *et al.* (2000)	adult	video	witness	–	–	>	>	–	>	*	>	–	*	*	–	>	>	–	–	*	–	*	>
Vrij, Edward *et al.* (2001a)	adult	video	witness	–	–	>	–	–	–	–	*	*	*	*	*	>	–	–	–	*	*	*	*
Vrij, Edward *et al.* (2001c)	adult	video	witness	–	>	>	–	–	–	*	–	–	*	*	>	>	>	–	–	*	–	*	>
Vrij & Heaven (1999)	adult	video	witness	*	*	–	*	*	*	*	*	*	*	*	*	*	*	*	>	*	*	*	*
Vrij, Kneller *et al.* (2000)[5]	adult	active	witness	–	–	>	–	–	*	–	>	–	*	*	*	>	>	–	–	*	*	*	>
Vrij & Mann (2006)	adult	active	suspect	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	>
Vrij, Mann *et al.* (2007)	adult	active	suspect	–	–	–	>	<	>	–	>	–	*	–	–	–	–	>	–	–	–	*	*
Vrij, Mann *et al.* (in press, normal order)	adult	active	suspect	*	–	*	–	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Vrij, Mann *et al.* (in press, reverse order)	adult	active	suspect	*	>	*	>	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Winkel & Vrij (1995)	8–9	video	witness	>	>	>	>	>	*	*	>	–	*	>	*	*	–	–	–	*	*	*	>
Total (support: total number of studies ratio)				11/27	12/24	22/29	16/26	12/27	15/25	5/22	11/25	7/23	1/10	4/14	8/24	7/21	8/26	9/23	4/20	0/10	2/10	1/3	16/20
Total support in percentages				41	50	76	62	44	60	23	44	30	10	29	33	33	31	39	20	0	20	33	80

Explanation of the signs: > verbal characteristic occurs more frequently in truthful than in deceptive statements; < verbal characteristic occurs more frequently in deceptive than in truthful statements; – no relationship between the verbal characteristic and truth telling/lying; * verbal characteristic was not investigated; na = participants participated in an activity, but they were neither a victim nor a suspect
Notes: [1]Uninformed truth tellers and liars only; [2]First interview only; [3]Single and repeated interviews combined; [4]Real and fabricated memories comparison; [5]Uninformed liars only

CHAPTER 9

Reality Monitoring

This chapter discusses a verbal veracity assessment tool called Reality Monitoring (RM). To my knowledge, professional lie catchers do not use this tool. However, it has attracted the attention of scientists worldwide, and to date researchers from Canada, Finland, France, Germany, Spain, Sweden, and the United Kingdom have published RM deception research (Masip, Sporer, Garrido, & Herrero, 2005; Sporer, 2004). A possible reason for its popularity amongst scholars is that the tool has a strong theoretical underpinning. RM in itself is not a veracity assessment tool, rather, it refers to the cognitive processes that are used to discriminate between perceived and imagined events (Johnson & Raye, 1981). Several scholars, however, believe that the RM concept is also relevant to deception and therefore use it in their attempts to detect deceit.

In this chapter I commence with outlining the RM concept. I will show that the original RM concept may be useful to detect false statements that are the result of people mistakenly believing that they have experienced an event. This sets RM aside from tools I have discussed in previous chapters. I already discussed that such false beliefs (memories) are difficult, and often impossible, to detect via analyses of nonverbal behaviour (Chapter 3) or via using SVA (Chapter 8). I then discuss how and under which circumstances the RM method could be used to detect deceit, followed by reviewing all RM research published in English that I am aware of. This chapter reveals that several RM criteria discriminate to some extent between truth tellers and liars and using this method can often lead to detecting truths and lies above the level of chance. I finally discuss how professional lie catchers could use RM analyses to assess the veracity of statements.

INSIGHT INTO PEOPLE'S MEMORY

RM differentiates between memory characteristics of actually experienced events and imagined events. The core of RM is that memories based on real experiences differ in quality from memories based on fiction. In 1981 Marcia Johnson and Carol Raye published their seminal paper about memory characteristics (see also Johnson, Hashtroudi & Lindsay, 1993, and Johnson & Raye, 1998). They argued that memories of real experiences are obtained through perceptual processes. They are therefore likely to contain *sensory information*: details of smell, taste or touch, visual details, and details of sound; *contextual information* which consists of spatial details (details about where the event took place, and how objects and people were situated in relation to each other), and temporal details (details about the time order of events, and the duration of events); and *affective information:* details about people's feelings throughout the event. These memories are usually clear, sharp, and vivid. By contrast, memories about imagined events are derived from an internal source and are therefore likely to contain *cognitive operations*, such as thoughts and reasonings ("I must have had my coat on, as it was very cold that night"). These are usually more vague and less concrete.

People often try to determine whether they have actually experienced an event they have in mind, or whether their memory is based on imagination. The processes by which a person attributes a memory to an actual experience (external source) or imagination (internal source) is called *Reality Monitoring*. We normally use the above-mentioned cues to discriminate between memories about actual experiences and imagined experiences. For example, when our memory about an event is vivid and contains a lot of sensory information, we are inclined to believe that we did experience the event, but we are more likely to assume that we have imagined the event if our memories are dim and contain lots of thoughts and reasoning (Johnson, Foley, Suengas, & Raye, 1988).

Let me give a real-life experience to clarify this concept. Once I was asked who else was in the car when we drove from A to B. When I mentioned three names, I was confronted with disbelief: "Are you sure?" I then started thinking. The driver was unfamiliar with the area and I could clearly remember the conversation between the driver and the person next to him, who knew the area well, about the directions the driver had to take. I could also visualise the gestures the person next to the driver made when giving directions. But who was sitting next to me? I assumed it was the driver's girlfriend, because how did she get home otherwise? However, I realised that I had no clear memory of her sitting next to me; rather, I used a cognitive operation (e.g., she is the girlfriend of the driver thus must have been in the car) to explain her presence in the car. I thus replied that, thinking about it, I was certain about the driver and the person sitting next to the driver but no longer certain about the presence of the driver's girlfriend sitting next to me. Indeed, it transpired that she was not in the car with me at that time.

To test people's memory for events, Johnson *et al.* (1988) have developed a 39-item Memory Characteristic Questionnaire (MCQ). This questionnaire measures the clarity and vividness of a person's memory for a particular event, as well as the memory of sensory, contextual and affective details or cognitive reasoning regarding the event. In one of their experiments, participants were instructed to rate the quality of their memory for specific events by completing the MCQ (Johnson *et al.*, 1988). Results showed that the quality of participants' memories did indeed differ between experienced and imagined events in the manner predicted by RM theory, however, these differences decreased as a function of time. That is, they found larger differences for adults' experienced and imagined recent memories than for their experienced and imagined childhood memories. It may well be that if people talk or think about an event, external memories (memories of experienced events) become more internal and internal memories (memories about imagined events) become more external.

One explanation of why external memories become more internal over time is that people use cognitive operations to facilitate and enhance later memory for experienced events (Roediger, 1996). For example, someone who drove fast through a foreign country to reach his holiday destination may try to remember the velocity with which he made his journey in two different ways. First, he could remember that he had frequently looked at his speedometer to check how fast he was driving. Alternatively, he could remember his speed by logical reasoning and by deducing that he must have driven fast because he drove on the motorway. The latter alternative, in which a cognitive operation is included, is an easier way of remembering that he drove fast than the first alternative. If this person is asked a couple of years later whether he drove fast through that foreign country it is therefore more likely that he will recall the information for his answer by remembering that he drove on the motorway than by remembering having regularly checked his speedometer. As a result, the person's memory about this experienced event will contain a cognitive operation. In contrast, imagined memories can become more external because people try to visualise what might have happened (Manzanero & Diges, 1996). As a result, their memories about imagined events will eventually become more vivid and concrete and may contain sensory, contextual, and affective details.

It also has been found that under some circumstances children do not differentiate between fact and fantasy as clearly as adults do (Lindsay, 2002). Children perform as well as adults when they are asked which of two other persons have done something, but perform worse than adults when they are asked which things they have actually done themselves and which things they have merely imagined having done. A number of factors may contribute to this (Ceci & Bruck, 1995; Lindsay, 2002; Lindsay & Johnson; 1987). For example, children may have a richer imagination than adults, and may therefore be better than adults at imagining themselves performing acts. Their memories for fact and fantasy will then become more similar. Also, adults have more efficient strategies to check whether their memory is accurate. Suppose you were with your family at a Mediterranean beach resort during the Christmas holiday. You talked with your family about going for a swim several times but you never went because it was too cold for a swim in the sea. A couple of years later a friend visits your family and asks whether you went swimming in the sea during that particular Christmas holiday at that beach resort. Adults may be less confused than children about this issue, because they realise that they would never have gone for a swim in the sea during the winter.

DETECTING FALSE MEMORIES

With the MCQ people rate the quality of their own memory. Is it possible for observers to rate the quality of someone else's memory, and could they thereby distinguish between real and imagined (false) memories? For observers to get access to other people's memories, people need to verbalise their memories. That complicates matters according to Johnson (1988). She believes that how people describe their memories of events differs from how they actually remember these events. People have a tendency to make their stories sound interesting and coherent. If necessary, they will fill gaps in their memory by including some information that they do not actually remember but that they think makes sense and is probably true (e.g., when you know that someone always wears a blue scarf, you may include in the description of a particular event that the person was wearing a blue scarf, although in fact you can no longer actually remember this detail). This tendency to fill in gaps particularly happens with imagined events, as they are less clear and vivid. As a result, differences between experienced and imagined events become smaller when people are asked to describe their memories (Johnson, 1988).

Schooler, Gerhard, and Loftus (1986) were perhaps the first researchers to test observers' ability to discriminate between real and false memories. Some participants saw a slide involving a yield sign, whereas another group did not see this sign but merely had its existence suggested. Many participants in both groups later reported to have seen the sign and gave verbal descriptions of the sign. This implies that some participants now falsely believed (e.g., imagined) having seen the sign. Observers, some of whom were trained to identify RM criteria, were then given transcripts of these verbal descriptions that included the sign, and asked to classify for each transcript whether the person had either imagined the yield sign or had really seen it. Trained observers were more accurate in detecting the imagined descriptions (60% accuracy rate) than untrained observers (50% accuracy rate), but trained observers were no more accurate in detecting the real descriptions (63%) than untrained observers (63%) (Schooler *et al.*, 1986, Experiment 4). Trained observers were thus to some extent able to distinguish between real and imagined descriptions, but their performance was not that impressive (62% overall accuracy, 12% above flipping a coin).

In another study, Porter, Yuille, and Lehman (1999) succeeded in getting their participants to believe they had experienced certain events (e.g., being sent to an emergency room) that they in fact never had experienced, via guided imagery. (See Chapter 2, Box 2.1, for more details about this study.) Participants gave verbal descriptions of these false memories and of events they had truly experienced. Both types of accounts were transcribed and rated for the presence of sensory components, but no differences were found. However, when participants were asked to rate their own memory of the events, they gave higher vividness ratings (an RM criterion) for their real memories than for their implanted memories, and this was a strong effect, $d = 1.20$.

In summary, the findings to date give only a weak indication that observers can distinguish between true and false memories by employing the RM tool. However, research in this area is scarce, and the issue has not yet fully been explored. For example, the descriptions of the yield sign used in Schooler *et al.*'s (1986) study were very short (20 words), and Porter and his colleagues examined some but not all RM criteria. I believe it is important to further investigate this issue of detecting imagined memories, particularly because it is often impossible to distinguish between real and imagined memories by looking at someone's behaviour or conducting SVA analyses. It is encouraging that people themselves appear able to discriminate to some extent between their own real and false memories by using the RM questionnaire (Porter *et al.*, 1999). However, there are restrictions. The RM imagined memories tool does not seem to work with young children or with adults when they refer to memories about events that happened a long time ago.

REALITY MONITORING AND DECEPTION

From 1990 onwards scientists have examined whether RM analyses can be used to discriminate between truths and lies. The assumption those scientists make is that truths are recollections of experienced events whereas lies are recollections of imagined events. Obviously not all lies are descriptions of events that a person did not experience. Many lies are not about events, but are about people's feelings, opinions, or attitudes (Chapter 2). However, one could argue that professional lie catchers are often interested in other types of lies, for example when people lie about their actions or whereabouts. Even when people lie about their actions and whereabouts they can sometimes describe events that they have actually experienced, but change the time at which they experienced them. For example, a burglar who denies having committed a burglary last night can claim that he went to the gym instead. He can then describe a truthful recollection of a visit to the gym (but on another occasion) and the only aspect he has to lie about is at what time or day he went there (*embedded lie*, Chapter 8).

If deceptive statements are based on actions and whereabouts that the liar did not experience but only pretend to have happened, do they then differ in RM criteria from truthful statements about experienced events? Researchers have addressed this question and I will describe the findings they have reported. The typical procedure of an RM deception experiment is that truth tellers and liars are interviewed, and these interviews are taped and transcribed. RM experts check for the presence of RM criteria in these transcripts. A standardised set of RM criteria has not been developed and so different researchers use different RM criteria. In this chapter I will use the set of eight criteria reported by Sporer (1997) because this was the first set of criteria published in English. The criteria are listed in Table 9.1. There is an overlap between some of the RM criteria and some of the CBCA criteria (Chapter 8), and I have indicated these similarities in the text below. Criteria 1 to 7 are expected to occur more often in truthful statements, whereas Criterion 8 is expected to occur more often in deceptive statements. Different

Table 9.1 Reality Monitoring criteria

1. Clarity
2. Perceptual information
3. Spatial information
4. Temporal information
5. Affect
6. Reconstructability of the story
7. Realism
8. Cognitive operations

researchers also use different ways of rating the criteria. Some scholars rate the frequency of occurrence of the criteria in the transcript, others use rating scales, and yet others use both frequency and rating scales.

Criterion 1: Clarity

This criterion refers to the clarity and vividness of the statement. This criterion is present if the statement is clear, sharp, and vivid (instead of dim and vague).

Criterion 2: Perceptual Information

This criterion refers to the presence of sensory information in a statement. Does the statement include sensorial experiences such as sounds ("He really shouted at me"), smells ("It had a smell of rotten fish"), tastes ("The chips were very salty"), physical sensations ("It really hurt"), and visual details ("I saw the nurse entering the ward")? The sound detail in RM is different from CBCA Criterion 6 *reproduction of conversation*. CBCA Criterion 6 is more restricted and only includes speech that is reported in its original form. Thus the two phrases "He said: Are you OK, you look so pale?" and "He asked how I felt" both count as a sound detail in RM, whereas the latter phrase does not count as a conversation detail in CBCA because it is not verbatim.

Criterion 3: Spatial Information

This criterion refers to information about locations ("It happened in a park") or the spatial arrangement of people and/or objects ("The man was sitting to the left of his wife"). This criterion is related to CBCA Criterion 4 *contextual embeddings*.

Criterion 4: Temporal Information

This criterion refers to information about when the event happened ("It was early in the morning") or explicitly describes a sequence of events ("When the visitor heard all that noise, he became nervous and left"). This criterion is also related to CBCA Criterion 4 *contextual embeddings*.

Criterion 5: Affect

This criterion refers to information that describes how the participant felt during the event ("I was very scared"). It is related to CBCA Criterion 12 accounts of subjective mental state, although CBCA Criterion 12 not only includes affect but also thoughts (Chapter 8).

Criterion 6: Reconstructability of the Story

This criterion examines whether it is possible to reconstruct the event on the basis of the information given. As such it is related to CBCA Criteria 1 to 3, *logical structure*, *unstructured production*, and *number of details*.

Criterion 7: Realism

This criterion examines whether the story is plausible, realistic, and makes sense. It is related to CBCA Criterion 1 *logical structure*. A difference, however, with CBCA is that this RM criterion takes plausibility into account whereas CBCA Criterion 1 does not.

Criterion 8: Cognitive Operations

This criterion refers to descriptions of inferences made by the participant at the time of the event ("It appeared to me that she didn't know the layout of the building"). In our own research, however, we use a wider definition and also include descriptions of inferences made by the participant when describing the event at a later time. For example, "She looked smart" is an inference of visual details and counts as a cognitive operation in our coding system.

INTER-RATER RELIABILITY SCORES

Inter-rater reliability refers to whether two RM raters who code the same statement independently from each other will obtain the same results. Research typically reveals satisfactory agreement between raters in RM coding (see Sporer, 2004, for RM inter-rater agreement scores). The inter-rater agreement typically found is comparable to that in CBCA coding. My own experience with both RM and CBCA coding is that it is easier and less time consuming to teach coders the RM tool than the CBCA tool. I think there are two reasons for this (see also Sporer, 1997). First, the RM tool contains fewer criteria than the CBCA tool, so the RM evaluator has less to learn. Second, there is less room for (mis)interpretation in coding RM criteria than in coding CBCA criteria. For example, raters have no difficulty in distinguishing between *visual*, *sound*, *spatial*, or *temporal* details (RM criteria), but experience some difficulty in determining whether a detail is *unusual*, *superfluous*, or *unexpected* (CBCA criteria).

REALITY MONITORING AS A LIE DETECTION TOOL

Appendix 9.1 provides an overview of the RM deception studies that I am aware of that have been published in English. Several studies have been published in German, French and Spanish, which restrict their accessibility to many readers and hence I have left them out, but the German and Spanish scholars Siegfried Sporer and Jaume Masip review those studies (in English) elsewhere (Masip, Sporer *et al.*, 2005; Sporer, 2004).

As I did with the CBCA studies, I have also included conference presentations which have been discussed elsewhere (Masip, Sporer *et al.*, 2005; Pezdek & Taylor, 2000; Sporer, 2004). Many RM studies have been published as book chapters in the series of the European Association of Psychology and Law. Those chapters are not as thoroughly reviewed as journal articles, so their quality is not guaranteed. Many of these papers, as well as the conference papers, do not give detailed information about how the RM criteria were operationalised. This is problematic because RM is not a standardised tool and different researchers seem to code the criteria somewhat differently. Many of those papers do not give information about the reliability of their coding systems either. I have the impression that only one evaluator coded the statements in many of those studies, and hence we do not know whether the coding is reliable. I have therefore included in Appendix 9.1 whether the paper has been published in a peer reviewed journal or not. I already mentioned that the RM tool may not work with young children. I have therefore also mentioned in Appendix 9.1 whether the participants were adults or children, and if the latter then their age. None of the studies examined memories of events that happened a long time ago, so the assumption that RM may not work when people describe events from the past cannot be examined on the basis of the research listed in Appendix 9.1.

Almost all studies were laboratory studies in which participants (mostly adults) were asked to describe something that they had or had not actually experienced. For example, truth tellers in Höfer, Akehurst, and Metzger's (1996) experiment took part in a photo-session and were interviewed afterwards about what happened. Liars did not take part in the photo-session but were told what the truth tellers had experienced. In the subsequent interview they had to pretend that they actually had their photo taken. Roberts, Lamb, Zale, and Randall (1998) and Rassin and van der Sleen (2005) carried out field studies. Kim Roberts and her colleagues examined statements of 26 alleged sexually abused children whose cases were confirmed ($n = 10$) or unconfirmed ($n = 16$). Cases were considered *confirmed* if the suspect confessed to the alleged incidents before plea bargaining, and/or if there was medical or physical evidence. The allegations were considered *unconfirmed* if there was a lack of substantiating evidence, the suspect persistently denied the allegations, and/or polygraph tests showed a truthful outcome regarding their denial of the allegation. I have already made clear in previous chapters that those factors are not independent case facts, which means that the ground truth was not established in this study. Rassin and van der Sleen's (2005) study has already been described in Chapter 8. They examined 27 true and 14 false allegations of sexual offences, but did not report the age of the alleged victims. An allegation was considered true if the allegation resulted in a conviction of the accused, whereas an allegation was considered false if the victims themselves were convicted for making a false allegation. Convictions are not independent case facts (Chapter 8), so the ground truth has not been established in this study either.

If a criterion occurred more frequently in true statements than in false statements, it is indicated with a ">" sign in Appendix 9.1; a "<" sign means that a criterion was less often present in true than in false accounts; if no difference was found between true and false statements, a "–" sign is reported. Empty cells mean that the criterion was not investigated in that particular study. According to RM theory, all criteria should occur more often in truthful statements, except the criterion *cognitive operations,* which should occur most frequently in deceptive statements.

Appendix 9.1 generally shows a mixed pattern. *Clarity* is not always a diagnostic sign of truthfulness, but when a difference between truth tellers and liars does emerge, it is always found that truthful statements are clearer than deceptive statements. Findings for *sensory information* are unclear, but how this criterion was operationalised and coded was not reported in many of the studies where it was examined. In journal articles where a distinction was made between *visual* and *sound* details a pattern emerged. Sound details in particular are more likely to be present in truthful accounts than in deceptive accounts. The findings about *contextual embeddings* are mixed, but, again, many of the studies where it was measured do not clarify how the criterion was defined and coded. Studies (mostly journal articles) that look at *spatial* and *temporal* details separately show a pattern. Those details are more likely to occur in truthful accounts than in deceptive accounts. There is one inconsistent finding, because Bond and Lee (2005) reported that truth tellers included fewer spatial details in their stories than liars. This study differs in one important aspect from all the other studies. They used an automatic computerised coding system rather than a manual coding system. I return to this issue later on in this chapter.

Many studies did not find a difference between truth tellers and liars in terms of reporting how they actually felt during the event (*affect*), and in the three studies where a difference was found, they contradicted each other. This criterion is related to CBCA Criterion 12. In CBCA research the criterion did yield more support than in RM research, although it is not amongst the criteria with the strongest support in CBCA in research (Appendix 8.1).

Several studies have revealed that truthful stories are easier to *reconstruct* than deceptive stories. This criterion is related to the three general CBCA criteria (Criteria 1–3) which also yielded considerable support in CBCA research. Moreover, several researchers have found that truthful stories sound more realistic than deceptive studies. This criterion is related to the verbal cue *plausibility*, discussed in Chapter 4. We saw in that chapter that plausible stories are more likely to be truthful than deceptive. Finally, most researchers did not find a difference between truth tellers and liars in terms of *cognitive operations*, and in one study it was found that, in contrast with what is predicted in RM theory, truth tellers included more cognitive operations in their stories than liars. We, however, have found several times that liars do indeed include more cognitive operations in their stories than truth tellers. I think this is related to how we coded this criterion. I do not know how others code this criterion due to a lack of detail about the coding systems in their reports, but we may have used a broader definition for cognitive operations than others. As I discussed above, we rate inferences that participants make both at the time of the event and at the time of recalling the event as cognitive operations. Thus, the sentence "She looked polite" contains a cognitive operation in our coding system, so do the sentences "She seemed quite clever" and "He looked tall for his age". Those inferences are all made at the time of recalling the event and are unlikely to be coded as cognitive operations by other researchers. Our broader definition of cognitive operations works well and appears to discriminate between truth tellers and liars.

As I already mentioned in previous chapters, examining single cues is the equivalent of examining Pinocchio's growing nose, which does not exist. Examining a cluster of cues (i.e., examining total RM scores) is a better strategy. Not many researchers have done this, but those who have calculated a total RM score typically find that truth tellers obtain a higher RM score than liars, as predicted by RM theory.[1] The exception is our study when we looked at the statements of 5–6 year olds. Total RM scores did not discriminate between truth telling and lying in 5–6 year olds but did distinguish between truths and lies told by participants who were 10 years of age or older. This finding supports the RM assumption that RM cannot be used with young children. Interestingly, CBCA scores did differentiate between 5–6 year old truth tellers and liars in the same study (Chapter 8), indicating that it is the sensitivity of the RM instrument rather than anything else that caused failure in the RM tool to detect the young children's lies.

Computerised Versus Manual RM Coding

Bond and Lee (2005) differed from other researchers in that they used an automatic, computerised, RM coding system. The advantages of employing a computerised coding system are obvious. Computerised coding could be much quicker than manual coding, and subjective differences between evaluators would cease to exist if all researchers were to use the same computer software. Bond and Lee used the Linguistic Inquiry and Word Count (LIWC) computer software program (Pennebaker, Francis, & Booth, 2001). LIWC counts the frequency of numerous words in texts and categorises these words. It contains many categories including the categories *senses*, *space*, *time*, and *cognitive mechanisms*, which are all related to RM criteria. In their experiment, Bond and Lee coded the RM criteria via LIWC. The results were mixed. Truth tellers obtained a higher score for sensory details than liars, but a lower score for spatial details than liars. The latter finding contradicts RM theory, and is the only erratic finding regarding spatial details published to date (Appendix 9.1).

Bond and Lee (2005) did not carry out manual RM coding on their data, so their study does not indicate how effective automatic RM coding

[1] Before calculating a total RM score, the cognitive operations criterion needs to be recoded, so that, similar to the other criteria that form part of the RM total scale, a high score indicates truth telling and a low score indicates lying.

(with LIWC) is compared to manual RM coding. We therefore examined this in our experiment (Vrij, Mann, Kristen, & Fisher, 2007). The results were unfavourable for LIWC coding. Although we obtained differences between truth tellers and liars in *sound*, *spatial*, and *temporal* details and in *cognitive operations* when we coded the transcripts manually (Appendix 9.1), no clear differences emerged via computerised LIWC coding. This is not surprising. Although the LIWC categories may resemble the RM categories, they are not developed on the basis of RM theory. This lack of theoretical foundation may cause error. For example, the LIWC cognitive mechanism category includes words such as *think*. Thus, the sentence "I think she had dark hair" would produce a hit in the LIWC cognitive mechanism category. By comparison, human RM coders do not count this as a cognitive operation. The problem with using automatic coding is that computer word counting systems ignore context, whereas the RM tool, as well as CBCA, require that the context is taken into account.

Issues Affecting RM Scores

RM is not a standardised test, and issues other than veracity will affect RM scores. Like CBCA scores, RM scores are dependent on age and throughout childhood and into adulthood RM scores will tend to get higher (Santtila, Roppola, & Niemi, 1999; Vrij, Akehurst, Soukara, & Bull, 2004a, b). Again, like CBCA scores, RM scores are related to personality traits. For example, people who are socially anxious obtain lower RM scores than individuals who are not socially anxious (Vrij, Akehurst *et al.*, 2004b). I already explained in Chapter 8 that it is difficult for evaluators to determine the exact impact of those factors when making their final veracity decision, and that they therefore tend to ignore such factors. Consequently, younger children and socially anxious people run the risk of not being believed when their statements are assessed with the RM tool, regardless of whether they are lying or not. Another issue is coaching. I discussed in Chapter 8 that those who come to know the CBCA tool could produce statements that appear credible to CBCA experts. No RM countermeasures study has been published to date, but I suspect that it is also possible for interviewees to produce statements that appear credible to RM evaluators if they learn how the RM tool works.

ACCURACY RATES IN REALITY MONITORING STUDIES

This section reviews the literature examining the extent to which RM analyses can distinguish truths from lies. In Table 9.2 I list all 10 studies

Table 9.2 Accuracy rates with the Reality Monitoring method, also in comparison with CBCA analyses

	Accuracy scores (RM)			Accuracy scores (CBCA)			Accuracy scores (RM+CBCA)		
Authors	Truth	Lie	Total	Truth	Lie	Total	Truth	Lie	Total
Granhag *et al.* (2006)	62	63	63			58			
Höfer *et al.* (1996)	61	70	66	70	73	71			
Santtila *et al.* (1999)	62	66	64	69	64	66[1]			
Sporer (1997)	75	68	71	70	60	65	83	75	79
Sporer & Sharman (2006)[2]	82	40	61						
Strömwall, Bengtsson *et al.* (2004)	53	73	63	44	64	54	70	68	69
Strömwall & Granhag (2005)[3]	82	83	82						
Vrij, Akehurst *et al.* (2004a)[4]	81	73	77	58	65	62			
Vrij, Akehurst *et al.* (2004b)[5]	88	61	74	50	69	60	88	61	74
Vrij, Edward *et al.* (2000)	71	64	67	65	80	73			
Total accuracy scores	**71.70**	**66.10**	**68.80**	**60.85**	**67.86**	**63.63**	**80.33**	**68.00**	**74.00**

Notes: [1]CBCA accuracy rates are from Santtila *et al.* (2000); [2]Veracity judgements were based on evaluating the JMCQ questionnaire; [3]Average scores of the two statements; [4]Adult participants only; [5]Lightly coached condition only.

that I am aware of that were published in English and reported accuracy rates.[2] Accuracy rates for truths vary considerably and range from a modest 53% to a very high 88%. The lie accuracy scores are more consistent and vary in most studies from 61% to 83%, however, the lie accuracy rate (40%) in Sporer and Sharman (2006) was exceptionally low. The average truth accuracy in those studies is 71.70%, and the average lie accuracy rate is 66.10%. The average total accuracy rate (truths and lies combined) is 68.80%. All of these rates are higher than could be expected by just flipping a coin.

In most of those studies researchers also carried out CBCA analyses, which makes a comparison of the two tools possible. The findings are inconclusive. In three studies, CBCA analyses resulted in superior total accuracy rates and in the other five studies the best accuracy rates were achieved with RM analyses. The average total accuracy rate for RM in those eight studies is slightly higher (68.13%) than the average total accuracy rate for CBCA (63.63%). Some researchers have examined whether the two tools combined produce higher accuracy rates than each individually. The combined tools yielded better results than the CBCA tool alone in all three studies, but were superior to the RM tool alone in only two of the three studies. In the third study, the combined tools were as accurate as the RM tool alone.

EVALUATION OF REALITY MONITORING AS A LIE DETECTION TOOL

The rationale behind Reality Monitoring is that memories of real experiences differ in quality from memories based on fiction, and that these differences in memories are reflected by differences in speech. The RM veracity assessment tool has restrictions and limitations. The restrictions are that it cannot be used with young children and probably cannot be employed when people talk about events that happened a long time ago. One limitation is that the RM criteria are poorly defined in deception research. Different researchers have used different definitions of the criteria, and the results regarding the *cognitive operations* criterion revealed how important a definition is. We found differences between truth tellers and liars in cognitive operations, but other researchers who operationalised cognitive operations differently did not. Another limitation is that RM is not a standardised test, and issues other than

[2] I left out the accuracy rates reported by Bond and Lee (2005). They did not include the spatial detail category in their RM model because the findings regarding this variable were the antithesis of that predicted on the basis of RM theory (Appendix 9.1). Bond and Lee's decision to leave the spatial category out artificially inflates the accuracy of their RM model.

veracity will impact upon RM scores. A similar problem occurred with CBCA. Like CBCA scores, RM scores are dependent on age and personality traits, and are probably vulnerable to countermeasures.

The research presented in this chapter shows that the RM tool can detect truths and lies above the level of chance. I therefore believe that it can be used in real life, and see two ways of employing the RM deception tool: (i) on its own, or (ii) in combination with CBCA. First, because truths and lies can be detected above the level of chance means that professional lie catchers can use the tool on its own. RM is not an ultimate lie detection test and the evidence gathered with this tool should not be allowed as evidence in criminal courts. However, RM analyses could be useful to form an indication about the veracity of statements.

Second, RM could be used in combination with CBCA (see also Sporer, 2004). This may have benefits. CBCA takes cognitive and motivational factors into account whereas RM is based on aspects of people's memory. Cognition, motivation, and memory are all relevant to deception and, as such, a combined tool may cover more aspects of deception than each tool individually. CBCA evaluators who use their CBCA tool to assess the veracity of alleged child victims in sexual abuse cases may wish to add the RM criterion *perceptual information* to their CBCA list. Pornographic films may increase children's knowledge about sexual acts. As a result, an inexperienced child may give a detailed account about a non-experienced sexual encounter after watching a pornographic film. However, in such an account details about smell and taste will be missing, as genuine experiences are required for such details. Such details in statements about sexual abuse may therefore be an indication that the statements are based upon real experiences. (Unless smell and taste were mentioned or indicated by people in the pornographic film.) Moreover, CBCA includes the general criterion *contextual embeddings* (Criterion 4), but RM makes a distinction between *spatial* and *temporal* details. Since both spatial and temporal details have been found to distinguish between truth tellers and liars, it would be interesting to see whether a more refined contextual embeddings criterion will enhance the accuracy of the CBCA tool.

In addition, the RM criterion *cognitive operations* could be added to the CBCA tool. The attractive aspect of this criterion is that its presence indicates lying. CBCA only contains criteria that indicate truth, and the presence of each criterion increases the likelihood that the story is truthful. However, the absence of a criterion does not imply lying; it only means that the statement lacks evidence of truth (Chapter 8). This may cause problems for the evaluator, because how should one interpret a very low CBCA score? The evaluator may be more confident that the statement is fabricated if it was to contain a lie symptom, because it sounds plausible that there is stronger indication of deceit if a statement both lacks the CBCA truth criteria and contains the RM deception criterion than if a statement just lacks the CBCA truth criteria. However, before adding RM criteria to the CBCA list, evaluators should realise that RM criteria do not appear to differentiate between truths and lies in young children, and thus be less useful in assessing young children.

Obviously, adding RM criteria to the CBCA list implies that there are even more criteria to take into consideration. A time-consuming tool is probably not warranted in cases where professional lie catchers only wish to get an indication about the veracity of statements without intending to present the findings as evidence in court. It is therefore worthwhile to examine whether some criteria of the original CBCA and RM tools could be left out of a combined CBCA/RM tool.

It would further be useful to know whether evaluators can detect the frequency of occurrence of criteria while listening to audiotapes rather than by reading transcripts, because this would make using the tool considerably easier. Our research sheds light on this issue (Vrij, Evans, Akehurst, & Mann, 2004). In a training session, a trainer gave five observers definitions and examples of four CBCA and five RM criteria. The observers and trainer then watched a videotaped interview and estimated the frequency of occurrence of each of the nine criteria in that interview. The ratings of the observers were then compared with those of the trainer and the trainer provided the observers with feedback about their performances. After 90 minutes of training in which five interviews were watched and discussed, all observers felt that they knew what was required to estimate the frequency of occurrence of the CBCA and RM criteria. They were then shown 52 interviews and watched each interview twice. After watching it the second time they estimated the frequency of occurrence of each of the nine CBCA and RM criteria. Their estimates were then compared with the ratings of CBCA and RM evaluators who had coded the transcripts. The findings revealed that the training had worked. First, the five coders agreed amongst each other about the frequency of occurrence of all criteria except the RM criterion *cognitive operations*; second, their estimates also correlated with the ratings of the RM and CBCA evaluators who coded the transcripts; and, third, differences obtained between truth tellers and liars on the basis of the transcript coding were also found on the basis of the observers' videotape coding.

Apart from using RM to detect truths and lies, the RM Memory Characteristic Questionnaire (MCQ) may have potential to distinguish truths from false memories, that is, descriptions of events that people think they have experienced whereas, in fact, they have merely imagined. However, not much real-life research has been carried out to date, and I believe that more research into this issue is needed before the use of the MCQ could be recommended for this purpose.

Appendix 9.1 Differences between truth tellers and liars in Reality Monitoring criteria

	Age	Peer reviewed	Clarity	Sensory information	Visual details	Sound details	Contextual embeddings	Spatial information	Temporal information	Affect	Reconstructability	Realism	Cognitive operations	Total
Alonso-Quecuty (1992) (immediate)	Adult	No		>			>							
Alonso-Quecuty (1992) (delayed)	Adult	No		<			<							
Alonso-Quecuty (1996)	Adult	No		–			>						–	
Alonso-Quecuty (1996)	Adult	No		–			–						–	
Alonso-Quecuty (1996)	8–10	No		>			<						–	
Alonso-Quecuty (1996)	8–10	No		–			–						–	
Alonso-Quecuty *et al.* (1997)	Adult	Yes		>			>						–	
Alonso-Quecuty *et al.* (1997)	9	Yes		>			>						–	
Bond & Lee (2005)	Adult	Yes		>				<	–	–			–	
Granhag *et al.* (2006)[1]	12–13	Yes	–		–	>		–	–		–	>	–	>
Hernandez–Fernaud & Alonso-Quecuty (1997)[2]	Adult	No		>			>						–	
Höfer *et al.* (1996)	Adult	No	>	–				–	>	–	>	>	>	
Manzanero & Digest (1996)[3]	Adult	Yes		<			<				>	>		
Porter *et al.* (1999)	Adult	Yes	–	–							–			

Rassin & van der Sleen (2005)	Unknown	Yes	>	–				>						
Roberts *et al.* (1998)	Child	No	–	>				>	>	–			–	
Santtila *et al.* (1999)	7–14	Yes	–	–				–	>	<	–	–	–	
Sporer (1997)	Adult	Yes	–	–				>	>	>	–	>	–	
Sporer & Sharman (2006)[4]	Adult	Yes	–	–				–	>			>		
Strömwall, Bengtsson *et al.* (2004)[5]	10–13	Yes	>		–	–		>	>	–	>	>	–	>
Strömwall & Granhag (2005)	11	Yes		–	–	>		–	>	>				
Vrij, Akehurst *et al.* (2004a)	5–15+ adult	Yes			>	>		>	>				<	
Vrij, Akehurst *et al.* (2004a)	5–6	Yes												–
Vrij, Akehurst *et al.* (2004a)	10–15+ adult	Yes												>
Vrij, Akehurst *et al.* (2004b)	6–15+ adult	Yes												>
Vrij, Edward *et al.* (2000)	Adult	Yes			>	>		>	>				<	>
Vrij, Mann *et al.* (in press, normal order)	Adult	Yes			–	–							–	
Vrij, Mann *et al.* (in press, reverse order)	Adult	Yes			–	>							<	
Vrij, Mann *et al.* (2007)[6]	Adult	Yes			–	>		>	>				<	>

Explanation of the signs: > verbal characteristic occurs more frequently in truthful than in deceptive statements; < verbal characteristic occurs more frequently in deceptive than in truthful statements; – no relationship between verbal characteristic and lying/truth telling; empty cells mean that the verbal characteristic was not investigated.

Notes: [1]Single and repeated recalls combined; [2]Cognitive interview and standard interview combined; [3]Spontaneous and planned interviews combined; [4]Other-ratings results. Coding was not based on transcripts but on the JMCQ questionnaire; [5]Single and repeated recalls combined; [6]Manual coding only, and accusatory interviews excluded.

CHAPTER 10

Scientific Content Analysis

This chapter discusses a verbal veracity assessment tool called Scientific Content Analysis (SCAN), which was developed by the former Israeli police lieutenant and polygraph examiner Avioam Sapir. Initially it was not my intention to include SCAN in this book. There is little research carried out with SCAN and none of the researchers who have examined the tool strongly recommends its use. However, I changed my mind during an investigative interviewing conference held in July 2006 at the University of Portsmouth (UK). At that conference I gave a seminar on lie detection to an audience that included criminal investigators from many different countries. When I asked which lie detection tool, if any, they mainly use, the most frequent answer was SCAN. The SCAN website http://www.lsiscan.com mentions that SCAN is being used by federal, law enforcement, and military agencies throughout Australia, Canada and the US, and that it is also used in many other countries including Belgium, Israel, Mexico, Singapore, South Africa, the Netherlands, and the United Kingdom. In other words, it appears to be used worldwide.

First, I will introduce the SCAN procedure. I will then discuss to what extent SCAN can discriminate between truth tellers and liars, followed by a discussion of its strengths and weaknesses.

THE SCAN PROCEDURE

The underlying assumption of SCAN is that a statement derived from memory of an actual experience differs in content and quality from a statement based on invention or fantasy (Smith, 2001). It is therefore thought that some SCAN criteria are more likely to occur in truthful statements than in deceptive statements, whereas other criteria are more likely to occur in deceptive statements than in truthful statements. Although SCAN predicts differences between truth tellers and liars, no theoretical rationale is provided as to why these differences would occur.

Written statements are used for SCAN analyses. It is important that the statement reflects the examinee's own words. Investigators can transcribe oral statements, but preferably, examinees write their statements themselves because this reduces the risk of the statement being contaminated by the views of the investigator. Examinees are requested to write down their activities during a certain period of time (e.g., "What did you do from the time you woke up until you went to bed?"). They should write their statements in enough detail so that a reader without background information about the activities of the examinee could fully grasp the statement. They can write this statement during the interview (Adams, 1996), or prior to the interview by filling out a VIEW questionnaire (Sapir, 1987/2000). The investigator is not present when the examinee fills out this questionnaire, which avoids the risk of the investigator influencing the statement. Examinees can complete the questionnaire at their preferred time and location and return it by mail or fax. The statement needs to be handwritten (Sapir, 1987/2000) because, as I will explain below, one SCAN criterion examines the number of corrections made by the examinee. SCAN can be applied to statements from suspects and witnesses and can be used with adults and children (Sapir, 1987/2000).

In terms of coding, an investigator checks whether a statement contains elements that are in alignment with the predictions of SCAN. If this is the case, the investigator could then decide whether the examinee is likely to be lying, or, alternatively, could further discuss in a subsequent interview with the examinee the elements of the statement that raised suspicion.

The list of SCAN criteria is extensive. I will discuss the criteria that are most emphasised in workshops on the technique (Driscoll, 1994) or are used in research (Smith, 2001). Some of the SCAN criteria are similar to criteria that appear on the CBCA list (Chapter 8), as I will indicate in the text below. Interestingly, this includes criteria that SCAN evaluators believe appear more often in *deceptive* statements, whereas

Table 10.1 Scientific Content Analysis criteria

1. Denial of allegations
2. Social introduction
3. Spontaneous corrections
4. Lack of conviction or memory
5. Structure of the statement
6. Emotions
7. Objective and subjective time
8. Out of sequence and extraneous information
9. Missing information
10. First person singular, past tense
11. Pronouns
12. Change in language

CBCA evaluators believe that they appear more often in *truthful* statements.[1]

Criterion 1: Denial of Allegations

This criterion refers to whether the examinee directly denies the allegation in the statement by stating "I did not....". It is thought that truthful suspects are more likely to include denials in their statements than deceptive suspects.

Criterion 2: Social Introduction

This criterion refers to how the persons described in the statement are introduced. While writing the statement the writer should take care that it is clear to a reader to whom the writer refers to. Social introductions should therefore be unambiguous (e.g., "My wife Lisa..."). When the writer is ambiguous and for example fails to introduce someone (e.g., "We went outside" without mentioning who "we" are) it is thought that something in the writer's mind prevented him or her from introducing the other person. Other forms of ambiguous introductions are referring

[1] I think that some of these conflicting predictions arise from different predictions concerning liars' strategies. The underlying assumption of CBCA is that liars wish to convince target persons that something took place which, in fact, did not happen. That is, pretending that sexual abuse took place, or, if CBCA can be used in situations other than sexual abuse, pretending that activities other than sexual abuse did occur ("I went to the gym last night, and did this, and this and this"). The underlying assumption of SCAN seems to be that liars merely concentrate upon concealing certain activities without putting much effort into convincing target persons that something else took place ("I watched television but I can't remember what was on").

to some people by pronoun ("he", "she", etc.) and others by their personal name; or suddenly omitting personal names at some places in the statement where they used these personal names in other parts of the statement. Ambiguity in social introductions may indicate that the writer is hiding something or may indicate tension between the people to whom the ambiguity refers.

Criterion 3: Spontaneous Corrections

This criterion refers to the presence of corrections in the statement, such as crossing out what has been written. Although explanations and additions are allowed, examinees are explicitly instructed not to cross anything out. A failure to follow this instruction could indicate deceit. This criterion is similar to CBCA Criterion 14 but CBCA experts believe that spontaneous corrections indicate truthfulness rather than deceit.

Criterion 4: Lack of Conviction or Memory

This criterion is present when the writer is vague about certain elements in the statement ("I believe..", "I think...", "kind of...") or when the writer reports that he or she cannot remember something. SCAN users interpret these phrases as suspicious. This criterion is similar to CBCA Criterion 15 but CBCA experts interpret lack of memory as a sign of truthfulness rather than deceit.

Criterion 5: Structure of the Statement

This criterion refers to the balance of the statement. It is thought that in a truthful statement the first 20% is used to describe activities leading up to the event, the next 50% to describe the actual event, and the final 30% to discuss what happened after the event. Thus, a 10-line statement is thought to comprise two lines to introduce the event, five lines to describe the event, and three lines about the aftermath. The more unbalanced a statement, the greater the probability that the statement is deceptive.

Criterion 6: Emotions

This criterion refers to whether there are emotions described in the statement. Truthful statements are more likely to contain descriptions of emotions than deceptive statements. This is identical to CBCA Criterion 12. However, unlike CBCA, in SCAN this criterion also refers to *where* the emotions are mentioned in the statement. It is thought that deceivers will mention emotions just *before* the climax of the story, whereas truth tellers are more likely to mention emotions *throughout* the story, but particularly *after* the climax of the story. Smith (2001) gives the example of a man suspected of insurance fraud related to injuries sustained at work. He claimed that he fell at work. In his statement he wrote "....as I was trying to climb up....I got very nervous. I was afraid I might fall...." The SCAN user indicated the location of the emotion within the statement (just before the climax of the story) and the man confessed to fraud.

Criterion 7: Objective and Subjective Time

This criterion refers to how different time periods are covered in a statement. Objective time refers to the actual duration of events described in the statement, whereas subjective time refers to the amount of words spent to describe these events. It is thought that in a truthful statement the objective and subjective time will correspond with each other, unlike in a deceptive statement. For example, if someone devotes five lines in a statement to describe a 30-minute period, and then three lines to describe a subsequent two-hour period, the objective and subjective time do not correspond and this may indicate deceit.

Criterion 8: Out of Sequence and Extraneous Information

This criterion examines whether the statement recounts the events in chronological order. A deviation of the chronological order may be deceptive. This criterion further examines whether there is extraneous information that does not seem relevant. If irrelevant information is provided, the SCAN expert should investigate why the examinee felt the need to include it. It is thought that examinees could include extraneous information to hide more important information. This criterion is a combination of two CBCA criteria: *unstructured production* (Criterion 2) and *superfluous details* (Criterion 9). CBCA experts, however, rate these criteria as signs of truthfulness rather than as signs of deceit.

Criterion 9: Missing Information

This criterion refers to phrases in the statement that indicate that some information has been left out. Examples are the use of words such as "sometime after", "finally", "later on", and "shortly thereafter". Smith (2001) gives the following example "She began hitting and kicking me, finally she hit me with the wine bottle". This sentence indicates that the writer does not wish to reveal what happened in the period between the hitting and kicking started and the hitting with the bottle occurred.

Criterion 10: First Person Singular, Past Tense

This criterion refers to the format in which a statement is written. It is thought that truthful statements are written in the first person singular, past tense because the writer describes an event that has taken place ("I saw the smoke coming out of the window"). Deviations from this norm should raise suspicion.

Sometimes, however, the use of present tense is expected, and the use of past tense can indicate deceit (Adams, 1996). Take for example statements that refer to missing persons. When children are missing, they are typically still alive in their parents' hearts and minds and the children should thus be referred to in the present tense (Adams, 1996). Deviation of this norm ("X loved playing with her dog and little sister") should raise suspicion.

Criterion 11: Pronouns

This criterion refers to the use of pronouns ("I", "my", "he", "his", "they", "their" etc.) in the statement. Pronouns signal commitment, responsibility, and possession. Omitting pronouns ("Left the house" rather than "I left the house") suggests reluctance of the writer to commit him/herself to the described action. The use of 'we" when "I" is appropriate suggests that the writer tries to absolve him/herself of personal responsibility. Leaving out pronouns that indicate possession ("my" etc.) suggests that the writer denies ownership. Smith (2001) gives as an example the following statement of a man who reported that his car was stolen: "Parked in . . .parking lot section G. Went shopping for half an hour. Came out it was gone". The statement shows an absence of pronouns, and thus indicates an absence of commitment, responsibility, and possession. The man later confessed that his car was not stolen.

The use of pronouns may also identify the relationship between two people. For example, if someone in a statement continually refers to "My wife and I" rather than "We" this may indicate some distance or tension between the writer and his wife.

Criterion 12: Change in Language

This criterion refers to the change of terminology or vocabulary in the statement. A change in language indicates that something has altered in the mind of the writer. For example, if a suspect refers in his statement to all conversations he had as "conversations" except one conversation which he describes as a "discussion", it is likely that he perceived this conversation differently from the other conversations. When a change of language is noticed in a statement a SCAN user should contemplate whether the sequence of events justify such a change. If the SCAN expert can think of a justification, the writer may be truthful; if the expert cannot think of a justification, the writer may be deceptive. For example, a man changed his language when describing his car. It was referred to as "the vehicle", "the car" and "the dark coloured car". It is difficult to justify this change in wording and the man confessed that this car did not actually exist (Smith, 2001).

SCAN AS A LIE DETECTION TOOL

Only three studies examining the SCAN technique have been published to date. The first study was a field study carried out by Lawrence Driscoll (1994). He analysed 30 written statements voluntarily given by suspects immediately prior to their polygraph tests. All statements came from one polygraph examiner. The examinees were asked to write down in their own words everything they felt would prove they were telling the truth. They were instructed to start "at the beginning" and to be as complete and descriptive as they felt necessary to tell their story. Each statement was analysed for the presence or absence of SCAN criteria but no information is given about who coded the statements. The SCAN coder (I assume there was only one) allocated 8 out of 11 (73%) truthful and 18 out of 19 (95%) deceptive statements correctly. When examining the individual criteria, Driscoll found that none of the criteria could perfectly distinguish truth tellers from liars. *Denials of allegations* (Criterion 1) was the best predictor for each group and on the basis of this criterion 77% of truth tellers and 88% of liars could be correctly classified.

Despite these rather high accuracy scores, Driscoll does not fully endorse using SCAN. He found the relatively large group of truth tellers who were judged to be deceptive (27%) problematic and stated that "any action based solely on the analysis of the written statement should be reconsidered" (p. 86). In addition, Driscoll acknowledged an important limitation of the study: the ground truth was uncertain for all of the statements that were analysed. In other words, for each suspect it is unknown whether he or she was actually telling the truth or lying. Driscoll concludes that his findings therefore should be considered with caution. I would like to go further, and believe that the absence of a ground truth regarding every single case that was examined means that the accuracy rates cannot be trusted at all.

Ground truth was not a problem in the second study that examined some of the SCAN criteria (Porter & Yuille, 1996). In this laboratory study participants did or did not commit a mock theft and were interviewed about this alleged crime. However, in this experiment participants gave oral statements that were subsequently transcribed. This deviates from the preferred SCAN procedure of asking examinees to write down their statements. Porter and Yuile (1996) examined three SCAN criteria: *structure of the statement* (Criterion 5); *missing information* (Criterion 9); and *first person singular, past tense* (Criterion 10). Truthful and deceptive statements did not differ from each other regarding any of these criteria.

The third and final published SCAN study was a field study carried out by Nicky Smith (2001) for the British Home Office. She asked five groups of assessors to judge 27 statements. Three groups were SCAN users with different levels of experience, a fourth group were experienced detectives not trained in SCAN, and the fifth group were newly recruited officers also not trained in SCAN. The three groups of SCAN users could give truthful, deceptive or inconclusive verdicts, and correctly classified at least 80% of the truths and 75% of the lies. This sounds impressive but the group of experienced detectives untrained in SCAN obtained accuracy rates that did not significantly differ from the accuracy rates of SCAN users. In other words, knowledge of SCAN did not lead to a superior ability in distinguishing truths from lies. Moreover, as in Driscoll's study, the ground truth for all cases was uncertain, and this again means that the accuracy rates cannot be trusted.

Smith (2001) also examined how effectively the individual criteria could distinguish truth tellers from liars, and whether SCAN experts use some criteria more frequently than others. Three criteria revealed differences between truth tellers and liars: *denial of allegations* (Criterion 1); *missing information* (Criterion 9); and *tense change* (Criterion 10), and the differences were in accordance with the SCAN predictions (fewer denials, more missing information, and more tense changes in the deceptive statements).

Smith (2001) further found that the criteria most frequently used were *improper use of pronouns* (Criterion 11), *lack of conviction or memory* (Criterion 4), and *change in language* (Criterion 12). However, Smith also noticed that different experts used different SCAN criteria to justify their decision of whether or not a statement was deceptive. In other words, there was a lack of consistency in the application of SCAN amongst SCAN users. This lack of consistency also becomes apparent on the SCAN website. On that website Sapir discusses numerous examples of how his analysis of statements could detect deception. This includes statements made by Bill Clinton, Osama Bin Laden, President George Bush, former US Secretary of State Colin Powell, and a leader of Hamas. Sapir goes on to analyse a statement released by North Korea about its nuclear test on 9 October 2006. However, in the different examples, Sapir gives different reasons as to why each of the statements may indicate deceit.

In summary, Smith found that SCAN users were no better lie detectors than experienced detectives untrained in SCAN, and that there is a lack of consistency amongst SCAN users in applying the method. Based on her findings Smith (2001, pp. 38/39) concluded that her study "has highlighted the need for caution before its widespread introduction is considered".

Several SCAN criteria are also part of the CBCA method, and in CBCA research it has been tested whether these overlapping criteria differentiate truth tellers from liars. CBCA research, as summarised in Appendix 8.1, revealed support for the assumption that truth tellers report more emotions in their statements than liars (SCAN Criterion 6 and CBCA Criterion 12). However, no support was found for the other overlapping criteria. SCAN predicts that liars make more *spontaneous corrections* (SCAN Criterion 3), admit more *lack of memory* (SCAN Criterion 4), and report more *out of sequence and extraneous information* (SCAN Criterion 8). CBCA research showed that in many studies truthful and deceptive statements did not differ from each other regarding these criteria, and, when differences did emerge, the opposite pattern was found from what is predicted by SCAN: liars make *less* spontaneous corrections (CBCA Criterion 14), show *less* lack of memory (CBCA Criterion 15), and report *less* out of sequence (CBCA Criterion 2) and extraneous information (CBCA Criterion 9) than truth tellers.

EVALUATION OF THE SCAN TECHNIQUE

The underlying assumption of SCAN is that a statement derived from memory of an actual experience differs in content and quality from a statement based on invention or fantasy. Therefore, it is assumed that some criteria are more likely to appear in truthful statements than in deceptive statements, whereas other criteria are more likely to appear in deceptive than in truthful statements. Problems associated with SCAN are a lack of empirical support and lack of standardisation.

Regarding the lack of empirical support, hardly any research has been carried out to test whether the method can successfully discriminate between truth tellers and liars. Two studies found that many truth tellers and liars could be correctly classified with the method, but in both studies the ground truth could not be established. This means that it was uncertain whether the examinees were actually telling the truth or lying, and the accuracy rates in those studies can therefore not be trusted. Other studies, mainly CBCA research, revealed that some of the SCAN criteria did not differentiate between truth tellers and liars in the way predicted by SCAN.

Regarding the lack of standardisation, different SCAN users concentrate on different criteria when analysing the same statement; and in different statements different criteria are emphasised as being indicators for deceit. A lack of standardisation implies that much depends on the subjective interpretation and skill of the individual SCAN user. As a result, individual differences will occur and some SCAN users will be better at distinguishing truths from lies than others. It would be useful to find out who are better and who are poorer lie detectors when using SCAN, but no method exists to discriminate between the two.

I believe that two factors contribute to the lack of standardisation in using SCAN. First, SCAN lacks theoretical underpinning. The underlying rationale of SCAN is that statements of truth tellers and liars will differ from each other but the method fails to explain why these differences would occur. A theoretical understanding of why differences would emerge between truth tellers and liars would provide different SCAN users with a common frame of mind about how to interpret the statements they evaluate, and this would promote consistency amongst them. Second, the list of SCAN criteria is a list of individual criteria rather than a set of cohesive criteria. In that respect, SCAN differs from the other verbal veracity assessment tools described in this book – BAI, CBCA, and RM (Chapters 7 to 9). In CBCA, the list of verbal criteria that is examined is cohesive. Therefore, when analysing a statement a total CBCA score based on the absence or presence of each of the individual criteria of the tool is calculated and a veracity assessment is based on this total score (Chapter 8). This does not mean that different CBCA evaluators will necessarily obtain the same total CBCA score when they evaluate the same statement, however, it does imply that different CBCA evaluators use the same assessment procedure when evaluating the statements. This principle also applies to BAI and RM. In contrast, SCAN contains criteria that are assessed individually. As a result, different evaluators may concentrate on different criteria when analysing the same statement; and the same evaluator may concentrate on different criteria when evaluating different statements.

Despite these criticisms, there are two elements in SCAN that I find attractive. First, I like the method of examining changes in wording within a statement. When discussing nonverbal cues to deception

(Chapter 3) I mentioned that examining changes in behaviour within an individual at different, comparable points in the interview is probably one of the most successful ways in correctly interpreting behaviour and detecting deceit. The SCAN approach of examining verbal changes within a statement is applying this method to verbal statements. I will discuss this comparison method in more detail in Chapter 15.

The second element I like is using SCAN as a tool to guide the subsequent interview (Adams, 1996). That is, instead of deciding whether or not an examinee has been deceptive, investigators can use the SCAN analysis to obtain insight into which elements of the statement they would like to further question the examinee. In this way, SCAN can be seen as an interview guidance tool bringing structure to the interview. Although Adams (1996) introduces SCAN as an interview guidance tool rather than a lie detection tool, I am not sure that all SCAN investigators use it that way. My suspicion is that many investigators use it as a lie detection tool and will make up their minds about the veracity of an examinee merely on the basis of the SCAN outcomes. (Sapir does this on his website.) If police detectives who use SCAN believe that a suspect has been lying in his or her statement, they may submit the suspect to persuasive interrogation techniques meant to break down the suspect's resistance. As I explained in Chapter 6, it is important not to submit an innocent suspect to these persuasive interrogation techniques, as it may lead to a false confession. There is no guarantee that SCAN outcomes are a safeguard for innocent suspects.

In conclusion, although I see some positive elements in SCAN, the lack of evidence that SCAN actually works and the lack of standardisation of the tool do not justify, in my opinion, the use of SCAN as a lie detection method. Before I can recommend its use for lie detection purposes, research needs to demonstrate its effectiveness and that different SCAN users show overlap in their use of the method. In that respect it is interesting what Sapir says about the accuracy of his method. In his workbook about SCAN he writes: "Many students who studied the SCAN technique, and are applying it in their job, report that the SCAN technique is equivalent to polygraph, if not better than that" (Sapir, 1987/2000, p. 121). If one were to subject this quote to the SCAN technique, its vagueness in terms of accuracy rates is worrying because it suggests that Sapir does not know himself how accurate SCAN is.

CHAPTER 11

Physiological Lie Detection: The Concern Approach

This is the first of three chapters which address physiological lie detection. Throughout history it has been assumed that lying is accompanied by physiological activity within the liar's body. For example, in 1000 B.C. the Chinese used to force suspected liars to chew rice powder and then to spit it out. If the resultant powder was dry then the person was judged to have been lying (Kleinmuntz & Szucko, 1984). There was a physiological basis for this assumption, as fear is associated with decreased salivation and a dry mouth (Ford, 2006).

Nowadays other physiological measures are examined by scholars and practitioners, most frequently electrodermal activity (EDA, such as sweating from fingers), blood pressure, and respiration. They are typically measured simultaneously with an instrument called the polygraph (from two Greek words, *poly* = many, and *grapho* = to write). The polygraph is a scientific measuring device that can display, via ink pens on to charts or via a computer's visual display unit, a direct and valid representation of various types of bodily activity (Bull, 1988). The polygraph records changes in EDA, blood pressure, and respiration accurately, and is able to measure even very small differences by amplifying signals picked up from sensors attached to different parts of the body. In the typical use of the polygraph, two metal electrodes, each placed on a finger, measure EDA; a blood pressure cuff placed around the bicep measures changes in blood pressure; and pneumatic gauges around the person's chest and stomach measure changes in the depth and rate of breathing.

People sometimes call a polygraph a lie detector but this title is misleading. A polygraph does not detect lies, but only physiological activity that is assumed to accompany telling a lie. Polygraph examiners have no option other than to measure deception in this indirect way, because a pattern of physiological activity directly related to lying, akin to Pinocchio's growing nose, does not exist (Saxe, 1991).

The polygraph in the deception arena has more utilities than simply detecting deceit. The stereotypical belief in American culture is that a polygraph test is practically infallible (National Research Council, 2003),[1] and this results in three more applications for polygraph testing. First, it could work as a deterrent. For example, perhaps sex offenders on probation will not undertake certain activities because they fear that a polygraph examination will reveal them; or perhaps a potential spy will not apply for a job in a national security agency because he or she believes that a pre-employment polygraph test will reveal his or her spying intentions. Second, the prospect of having to undergo a polygraph test may elicit a confession. For example, a criminal might believe he or she is better off by admitting their criminal activities voluntarily than by being accused of both these activities and lying about it if they fail the polygraph test. Third, the knowledge that polygraph tests are used may give the general public the impression that they are protected against criminals and so may give them a sense of security.

Polygraph examiners use various tests to which they assume that truth tellers and liars will respond differently. One set of tests is built upon several psychological premises, including the premise that the heightened physiological responses displayed by examinees during key periods of the test are the result of increased concern. Eminent proponents of these concern-based tests include Charles Honts, David Raskin, and John Kircher.[2] Such tests, however, are controversial and criticised as I will discuss below. Eminent critics include Gershon Ben-Shakhar, John Furedy, William Iacono, David Lykken, and Leonard Saxe.[3] Initially, David Raskin and David Lykken engaged in prolonged

[1] The views in other countries are unknown and there may be cultural differences in this belief (Bull, 2006).

[2] Honts, 1991, 1995, 2004; Honts, Kircher, & Raskin, 1996, 2002; Honts & Perry, 1992; Raskin, 1982, 1986, 1988, 1989.

[3] Ben-Shakhar, 1991; Ben-Shakhar & Furedy, 1990; Ben-Shakhar, Lieblich, & Bar-Hillel, 1982; Furedy, 1991a, b, 1993, 1996a, b; Honts, Kircher, & Raskin, 1996; Iacono, 2000; Iacono & Lykken, 1997; Lykken, 1959, 1960, 1988, 1991, 1998; Saxe, 1991. See also Faigman, Kaye, Saks, & Sanders (1997, 2002), Gale (1988), and Kleiner (2002) for edited books in which both sides are represented.

controversy over the reliability and validity of various polygraph tests. They have come into conflict in the scientific literature, as expert witnesses in court and as possible opponents in the legal process against each other. The other proponents and critics of concern-based polygraph testing have joined this debate. Interestingly, the critics are not against polygraph testing per se. In contrast, they all endorse it, but propose the use of a test that, rather than being based on the premise of increased concern, is based on the assumption that liars will show orienting reflexes when confronted with crucial details of a crime (i.e., signs of recognition of the details).

I believe that the difference between the concern approach and the orienting reflex approach is fundamental enough to justify discussing polygraph tests based on these two approaches in two separate chapters. The concern-based polygraph tests are discussed in this chapter and the orienting reflex-based polygraph test will be discussed in Chapter 12.

Concern-based polygraph tests are carried out in various countries across the world, including Belgium, Canada, China, Israel, Mexico, Norway, Pakistan, Poland, Russia, Romania, Singapore, South Korea, Taiwan, Thailand, the Philippines, the United States, and Turkey (British Psychological Society, 2004; Honts, 2004; Lykken, 1998; Svebak, 2006). Many tests take place in criminal investigations, although the outcomes of these polygraph tests are typically not used as evidence in criminal courts.[4] In the US polygraph tests are also conducted in civil proceedings and the outcomes are sometimes presented in civil courts (Iacono, 2000; Raskin & Honts, 2002). Perhaps most concern-based polygraph tests are carried out in the US, where they are used for at least four purposes (Consigli, 2002; Gale, 1988; National Research Council, 2003): to assess the truthfulness of examinees in criminal investigations; for pre-employment screening in law enforcement and national security agencies; for screening existing employees, especially in security-sensitive occupations; and for testing sexual offenders who are in treatment, on parole or in prison.

I will discuss three concern-based polygraph tests that are currently used, and describe what critics have to say about these tests. I also discuss the laboratory and field studies that have examined the accuracy of these tests. These studies, almost exclusively carried out in criminal investigation settings using one particular test, the Comparison Question Test (CQT), indicate that the CQT is able to discriminate between truth tellers and liars above the level of chance, but the accuracy levels are

[4] Exceptions are New Mexico State and Belgium where the outcomes of the tests are used as evidence in criminal courts (Daniels, 2002; Van der Plas, 2006).

far from perfect. This is one of the reasons why, in my view, concern-based polygraph tests should not be allowed as evidence in criminal courts. I also discuss why polygraph examinations for the screening of applicants, employees, and sex offenders are more problematic than polygraph examinations with crime suspects.

Apart from EDA, blood pressure and respiration, other physiological measures are examined with regard to their relationship with deception, such as changes in blood flow in the brain (brain activity measured with functional magnetic resonance imaging (fMRI) scans), event-related potentials (P300 brain waves measured with EEGs), characteristics of the voice signal (measured with voice stress analysers), and changes in blood flow patterns around the eye (thermal imaging, measured with heat detecting cameras). I will also discuss these physiological measurements. fMRI lie detection will be discussed in Chapter 13. EEG-P300 lie detection is based on the orienting reflex and will therefore be discussed in Chapter 12. Voice stress analyses and thermal imaging are concern-based techniques and are discussed in the present chapter.

CONCERN-BASED POLYGRAPH TESTS

The Relevant–Irrelevant Test

The Relevant–Irrelevant Test (RIT) was the first widely used polygraph test. It was developed by Larson in 1932 and is based on the research of Marston (1917). In the RIT, two types of questions are asked: crime-relevant questions and crime-irrelevant questions. Crime-relevant questions are related to the crime under investigation, such as "Did you break into a black Mercedes last night?" All suspects, both innocent and guilty, will answer "No" to this question, otherwise they are admitting to having committed the crime. Crime-irrelevant questions are unrelated to the crime, and the examiner knows that the examinee will tell the truth when answering these questions. An example of a crime-irrelevant question is "Is it Tuesday today?" The examiner will then compare the physiological responses to both types of questions. The rationale behind the RIT is that the observed physiological responses are produced by detection anxiety (Raskin & Honts, 2002). Therefore, larger responses to the crime-relevant questions than to the crime-irrelevant questions are interpreted as signs of lying to the crime-relevant questions.

This rationale has been described as simplistic and naive (Podlesny & Raskin, 1977). The crime-irrelevant questions are meant to control for interpersonal differences, that is, differences in physiological responses naturally shown by different individuals. However, as I explained in Chapter 3, not only *interpersonal* differences, but also *intrapersonal* differences should be taken into account, which is the notion that the same person may respond differently to different questions due to the nature of those questions. Suppose a woman is attached to a polygraph to discover whether she stole money from her employer's office. She is innocent, but she realises that a large physiological response might imply that she will lose her job. Hence, the crime-relevant question "Did you steal the money?" has major consequences for her. Thus, it is quite possible that this crime-relevant question will result in a larger physiological response than a crime-irrelevant question about, say, the colour of her shirt.

A similar case was broadcast on Dutch television a while ago (in a Veronica programme entitled *Berg je voor Berg*). The cyclist and former world champion Gerrie Knetemann was attached to a polygraph with the intention of discovering whether Knetemann had ever used illegal drugs in his career to enhance his performance. Both crime-irrelevant questions such as "Are you wearing a watch at the moment?" and "Are you now in Amsterdam?" and a crime-relevant question "Did you ever use illegal performance-enhancing drugs in your cycling career?" were asked. Knetemann answered "No" to the crime-relevant question, and was subsequently accused of lying, because his physiological response to this question was stronger than his physiological responses to the crime-irrelevant questions. This conclusion was, of course, premature. Performance-enhancing drugs are a sensitive issue for professional athletes and questions about this matter will almost automatically increase their physiological activity, whether or not they have actually used such drugs. In other words, all competitive sportsmen, both those who use drugs and those who do not, are likely to show strong responses to such crime-relevant questions.

These examples make clear that the RIT is an inappropriate lie detection test. The obvious problem with it is that the crime-irrelevant questions do not provide an adequate control for the emotional impact the crime-relevant questions can have on examinees (Iacono, 2000). There is agreement amongst proponents and critics of concern-based polygraph testing that the RIT should not be used (Honts, 1991; Lykken, 1998; Raskin, 1986; Raskin & Honts, 2002; Saxe, 1994). Larson himself acknowledged the limitations of his technique and declared that he was not particularly happy with the importance that others gave to it. He said in 1961: "I originally hoped that instrumental lie detection would become a legitimate part of professional police science. It is little more than a racket. The lie detector, as used in many places, is nothing more than a psychological third-degree aimed at extorting confessions as the old physical beatings were. At times I'm sorry I ever had any part in its development" (cited in Lykken, 1998, pp. 28–29).

Although the RIT technique has been the dominant polygraph technique for many years, it is now used infrequently in criminal investigations (Raskin & Honts, 2002). However, it is probably still used in other settings, such as when insurance companies attempt to detect fraud in their clients' claims, as I will explain later on in this chapter.

Comparison Question Test

Nowadays the most frequently used polygraph test in criminal investigations is the Comparison Question Test (CQT), also referred to as the Control Question Test. The CQT was initially developed by Reid in 1947 and further conceptualised by Backster (1962, 1963) and Raskin (1979, 1982, 1986). The main difference between the CQT and the RIT is that in a CQT examination an attempt is made to control for intrapersonal differences by introducing comparison questions that are an adequate control for the emotional impact that the crime-relevant questions may have on examinees.

The typical CQT consists of five phases (Furedy, 1991b; Raskin & Honts, 2002). The first phase, *the pre-interview phase*, is designed to obtain information about the suspect and the crime. The examiner asks the examinee some basic biographical information, including information about physical and mental problems, and discusses the allegations against the examinee. The examinee is encouraged to freely describe the incident from his or her point of view. This phase gives the examiner insight into the examinee's version of what has happened and also helps to establish an atmosphere of professional objectivity and trust.

The examiner then places the polygraph transducers on the examinee and explains in brief the physiology that underlies the polygraph test. The examiner then carries out a *stimulation test* (Phase 2), which is intended to convince the examinee that the polygraph is able to detect every lie. It is crucial for a polygraph test that the examinee believes that the test is infallible. The notion that the polygraph test is 100% accurate may augment concern in the guilty suspect when answering the crime-relevant questions ("There is no way to beat this machine"), and may increase confidence in innocent people when answering the crime-relevant questions ("The machine works accurately and, as I am innocent, I will therefore be exonerated"). The opposite may occur when examinees do not trust the accuracy of the polygraph: it may make guilty suspects more confident ("Nothing is lost yet, there is still a chance that I will beat the polygraph") and may increase concern in innocent suspects ("I know I am innocent, but is this machine going to show that?") when answering the crime-relevant questions.

A card game is often used in the stimulation test. The examinee is asked to pick a playing card from a deck of cards, to make note of it, and then return it to the deck. The polygrapher then shows the examinee several cards and asks each time whether this is the card the examinee had just seen. The examinee is instructed to respond with a "No" to each card. As the examinee does so, the examiner evaluates the polygraph charts and will inform the examinee which card he or she had just seen. The polygrapher will very often be correct. Showing the correct card will almost automatically lead to a physical response in the examinee due to the orienting reflex, which I will explain in Chapter 12.

Examiners always run the risk of making the wrong decision, and identifying the wrong card. This could have catastrophic consequences because it would demonstrate that the test is not infallible after all. To avoid a possible error, examiners sometimes use tricks unbeknown to the examinee, for example by marking the correct card or by using a deck of cards that contains only one type of card (Bashore & Rapp, 1993). In the latter case the examiner does not actually show the cards to the examinee but calls out the supposed card names. Other examiners may not play card games but persuade the examinee of the infallibility of the test by having a well-appointed office with various framed diplomas and certificates adorning the walls (Bull, 1988).

After the stimulation test, the *formulating the question phase* (Phase 3) commences. In this phase the examiner formulates the questions that will be asked during the polygraph test and discusses these with the examinee. There are two reasons for discussing the questions with the examinee. First, the examinee must understand the questions so that no further discussion about the content of the questions will take place during the polygraph examination or afterwards. Second, the examinee must be willing to answer the questions with definitive a "Yes" or "No" rather than with "Yes, but..." or "It depends..." answers, because in the latter case it could be difficult to interpret the outcomes of the test. Table 11.1 provides an example of a typical CQT question sequence in a case where somebody is suspected of stealing a camera. I will discuss a CQT about theft as an example, but a CQT could be administered to enquire about any type of crime.

There are three types of question: neutral, relevant, and probable lie questions. *Neutral questions* are general questions such as "Do you live in the United States?" and "Is your name Rick?" Neutral questions are used as fillers, and the physiological responses to these questions are disregarded when the polygraph charts are scored.

Table 11.1 An example of a Comparison Question Test sequence

N1.	Do you live in the United States? "Yes"
PL1.	During the first 20 years of your life, did you ever take something that did not belong to you? "No"
R1.	Did you take that camera? "No"
N2.	Is your name Rick? "Yes"
PL2.	Prior to 1987, did you ever do something dishonest or illegal? "No"
R2.	Did you take that camera from the desk? "No"
N3.	Were you born in the month November? "Yes"
PL3.	Before age 21, did you ever lie to get out of trouble or to cause a problem for someone else? "No"
R3.	Did you participate in any way in the theft of that camera? "No"

Notes: N = neutral; PL = probable lie; R = relevant
Source: derived from Raskin, Kircher, Horowitz, & Honts (1989)

A sequence typically contains three relevant and three probable lie questions. *Relevant questions* are specific questions about the crime. For example: "Did you take that camera?" is a question that can be asked in the camera theft example. *Probable lie questions* serve as control questions. They deal with acts that are related to the crime under investigation, but do not refer to the crime in question. They are always general in nature, deliberately vague, and cover long periods of time in the history of the examinee. The examiner formulates a probable lie question for which in the examiner's view a "No" answer would be deceptive. The exact formulation of the question depends upon the type of crime and the examinee's circumstances, but a probable lie question in an examination regarding theft might be: "During the first 20 years of your life, did you ever take something that did not belong to you?" where the examiner believes that the examinee did indeed steal something before age 21 (which many people have done). Under normal circumstances, some examinees might admit to this wrongdoing. However, during a polygraph examination they are unlikely to do this because the examiner will make the examinee believe that admitting to such a theft would demonstrate that he or she is the type of person who would commit the crime in question and so will be considered guilty. Raskin (1989, p. 254) gives the following example of how this could be achieved: "Since this is a matter of theft, I need to ask you some general questions about yourself in order to assess your basic honesty and trustworthiness. I need to make sure that you have never done anything of a similar nature in the past and that you are not the type of person who would do something like stealing that *camera* and then would lie about it. . . So, if I ask you, *"During the first 20 years of your life, did you ever*

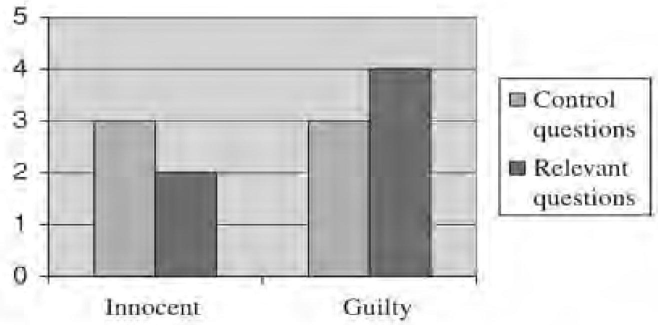

Figure 11.1 Physiological profile for innocent and guilty suspects (CQT)

take something that did not belong to you?" you could answer that "No", couldn't you?[5]

Most examinees answer "No" to these probable lie questions (Raskin & Honts, 2002). If, nevertheless, examinees admit to a past wrongdoing, the examiner will ask for an explanation. The examiner will then play down the examinee's admission (e.g., "Well that was when you were a child and didn't know better") and will reword the probable lie question (e.g., "Apart from what you have told me. . ." (O'Toole, Yuille, Patrick, & Iacono, 1994; Raskin & Honts, 2002)). Alternatively, after an admission to a probable lie question the examiner may respond with "OK, but did you take something that makes me think now that you stole the camera?"

The probable lie questions are designed to create a dilemma for the examinee (Raskin & Honts, 2002). On the one hand, the examinee will believe that admissions to probable lie questions will make the examiner think that they are dishonest and therefore guilty concerning the crime under investigation. On the other hand, the examinee will also believe that deceptive answers to probable lie questions will result in strong physiological reactions to such questions, and will lead the examiner to conclude that the examinee was deceptive with respect to the relevant issues concerning the crime in question and, going back to the example, may have been responsible for the theft of a camera. In fact, as I will now explain, the examiner will actually interpret strong physiological reactions to probable lie questions as a sign of *truthfulness* but will not inform the examinee about this.

Overall, probable lie questions and relevant questions are thought to evoke different patterns of physiological responses in guilty and innocent suspects, as is shown in Figure 11.1.

[5]The italic parts are slight modifications of Raskin's quote to bring the quote in line with the theft example I discuss.

Probable lie questions are thought to result in stronger physiological responses than the relevant questions in the *innocent suspect*. Since the examiner puts so much emphasis on the probable lie questions to which the examinee will be deceptive, and because the examinee knows he or she is answering truthfully to the relevant questions, the examinee will become more concerned about their answers to the probable lie questions than about their answers to the relevant questions. Consequently, they should react more strongly to the probable lie questions than to the relevant questions. However, the same probable lie questions are expected to elicit weaker physiological responses than the relevant questions in the *guilty suspect*. A guilty suspect responds deceptively to both types of question, which in principle should lead to similar physiological responses to both types of question. However, relevant questions represent the most immediate and serious threat to the examinee, which will make them more concerned about their lies to the relevant questions than about their lies to the probable lie questions. This will then result in stronger physiological responses to the relevant questions than to the probable lie questions.

Then the *test proper* (Phase 4) follows. It is important that examinees are not distracted during the polygraph test. Every distraction may result in a physiological response that will be detected by the polygraph and may influence the results. Therefore tests should preferably take place in a quiet room where outside noises and commotion cannot be heard, and the examiner and equipment are preferably located next to, rather than in front of, the examinee. Moreover, examinees are instructed to remain still as movements may lead to unwanted physiological responses. The necessity to sit still means that cooperation from the examinee is required during the polygraph test. Therefore, examinees only participate on a voluntary basis and can elect to withdraw at any time during the test.[6] The questions are presented at a rate of one every 25–35 seconds while physiological activity is recorded. The sequence is repeated at least three times.

The final fifth phase of the test, *the scoring phase*, is the interpretation of the polygraph charts, which can be done via a global or numerical approach. In the global approach, the examiner forms an impression of the examinee's physiological responses to the test. This impression is then combined in some unspecified manner with evaluations of the case facts (the examinee's criminal history, evidence in the case) and

[6] Examinees, however, are perhaps reluctant to do so, as this would make them appear more suspicious. Withdrawal may easily lead to reactions such as "If you are innocent, why don't you prove this in the polygraph test?"

the examinee's behaviour during the test in order to reach an overall decision about the truthfulness of the examinee.

Backster (1962, 1963) and Raskin (1979, 1982, 1986) advocated and introduced the numerical scoring approach. The numerical method is an attempt to systematically score the charts whilst minimising the influences of sources other than the polygraph charts in the decision making. In the numerical method comparisons are made between the reactions to the probable lie questions and subsequent relevant questions (PL1 compared to R1, PL2 compared to R2, and PL3 compared to R3). There are four possibilities. If there is no difference in physiological response, the score "O" will be assigned; if there is a noticeable difference the score "1" will be allocated; and strong and dramatic differences are indicated with a "2" and "3", respectively.

However, there is no standardised rule about what a "noticeable", "strong" and "dramatic" difference means, as this depends on the examinee. The same difference may be strong in one examinee but only noticeable in another. Figure 11.2 makes this point clear. Both examinees in Figure 11.2 showed a stronger response to the relevant question than to the probable lie question. The absolute difference in responses between the relevant and probable lie question is the same in both examinees. Yet, the relative difference is much larger in Examinee 2 than in Examinee 1, due to the fact that Examinee 2's overall physiological responses were smaller than Examinee 1's overall responses. The examiner will therefore assign a higher number to the difference in Examinee 2 than to the difference in Examinee 1.

Most assigned scores are 0 or 1, scores of 2 are less common, and scores of 3 are unusual (Raskin & Honts, 2002; Raskin, Kircher, Horowitz, & Honts, 1989). If the observed reaction is *stronger* to the relevant question than to the probable lie question, a negative score (−1, −2, or −3) is assigned. If the observed reaction is *weaker* to the

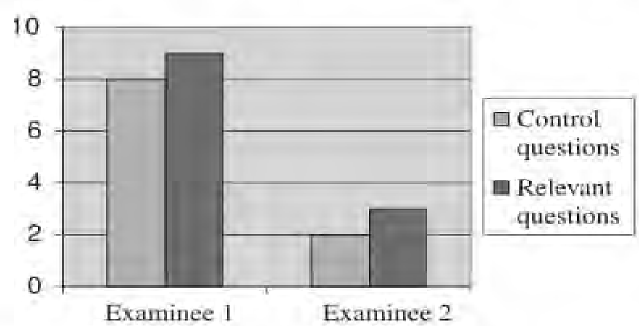

Figure 11.2 Individual differences in physiological responses

relevant question than to the probable lie question, a positive score (+1, +2, or +3) is allocated. Different scores are allocated to each physiological measure (EDA, blood pressure, and respiration). The scores are then summed over each of the three physiological measures and each of the three sequences to provide a total score. The outcome of the test is based on this total. Scores of +6 or higher indicate truthfulness, and scores of −6 or lower indicate deception; scores between +5 and −5 are inconclusive outcomes. The responses to the first probable lie question and relevant question are often disregarded, as examinees sometimes show disproportionately strong responses to the first question, due to unfamiliarity with the polygraph machine or nervousness about the investigation.

Box 11.1 ***An unofficial confession induced final phase***

An unofficial, sixth phase of the CQT involves informing the examinee directly after the test that he or she is lying, and asking him or her to consider why it is possible that the charts show this outcome (Lykken, 1998). To facilitate this thinking process the examiner will leave the room for a while. The aim of this phase is to obtain a confession, and this happened in a particular case in which after making the accusation the examiner left the room for a while to watch the examinee from another room via a one-way mirror (Lykken, 1998). The examinee, obviously upset, kept looking at the polygraph charts, then made up his mind and started to eat the charts, a total of six feet of paper, six inches wide. After he had finished this meal, the examiner returned as if nothing had happened, leaned his ear down to the polygraph and said: "What's that? He ate them?" The examinee responded by saying "My God, you mean the thing can talk, too?" and confessed to having committed the crime.

Criticism Concerning the Comparison Question Test

The Comparison Question Test is controversial and evokes criticism in its opponents. Iacono (2000) believes that there are three reasons that contribute to its controversy: (i) there is no consensus amongst scientists that there exists an adequate theoretical foundation for its application; (ii) the polygraph profession operates outside the scientific environment and is practised most by law enforcement officials trained at freestanding polygraph schools that are unrelated to universities;

and (iii) polygraph tests can have profound consequences for individuals subjected to them. In this section I will summarise the main points of criticism.[7]

A weak theoretical foundation

A main point of criticism is the theoretical foundation for the CQT. The National Research Council (2003, p. 213) describes it as "....quite weak, especially in terms of differential fear, arousal, or other emotional states that are triggered in response to relevant or comparison questions". Several theoretical accounts have been offered to justify the CQT assumptions – guilty examinees will show the strongest responses to the relevant questions, whereas innocent examinees will show the strongest responses to the probable lie questions – however, none of these accounts rule out the possibility that innocent examinees show larger physiological responses to relevant questions than to probable lie questions. In other words, the criticism is that the CQT does not adequately control for intrapersonal differences.

One theory put forward to justify the CQT premise is the *conflict theory* (Davis, 1961), which states that two incompatible reaction tendencies triggered at the same time produce a large physiological response. Lying to questions would trigger a conflict between "telling a lie" and "telling the truth" tendencies, and the more serious the lie (e.g., the more serious the crime), the stronger the conflict between the two tendencies becomes. However, a conflict between the examinee and examiner, for example, an expectation of being falsely accused, could also evoke conflict tendencies in innocent examinees (National Research Council, 2003).

Another theory used to justify the CQT premise, the *conditioned response theory* (Davis, 1961), states that relevant questions that specifically address a certain transgression elicit a strong recollection of that transgression in guilty examinees, which creates strong physiological responses. The less specific probable lie questions will not elicit such a strong recollection in guilty examinees. However, relevant questions may evoke all sorts of thought processes in innocent examinees, which could also create strong physiological responses (National Research Council, 2003). Consider the innocent suspect who is asked about the murder of his beloved wife. The mentioning of his wife may reawaken his strong feelings about her, which will be recorded on the polygraph charts.

[7] See also Ben-Shakhar, 2002; British Psychological Society, 1986, 2004; Fiedler, Schmid, & Stahl, 2002; Iacono, 2000; Lykken, 1998; National Research Council, 2003; Saxe & Ben-Shakhar, 1999.

The *threat of punishment theory* (Davis, 1961) states that if the examinee fears serious negative consequences of being caught in the lie, the threat of punishment when lying will result in large physiological responses. Since the negative consequences for guilty examinees are thought to be larger in lying to relevant questions than in lying to probable lie questions, the strongest physiological responses should be elicited by the relevant questions (National Research Council, 2003). However, innocent examinees may also associate serious negative consequences with not being believed when answering the relevant questions, and this could also result in strong physiological responses (National Research Council, 2003). What perhaps increases this problem is that polygraph tests are only carried out if no other conclusive evidence is available in the case. (Where there is conclusive evidence, a polygraph test is redundant.) This means for innocent persons that they have yet not been able to prove their innocence during the investigation prior to the polygraph test. For them, the outcome of the polygraph test really matters. A guilty verdict would mean that they are still not exonerated from the accusations or suspicions that have surrounded them. This could have serious negative consequences, such as continuing to be investigated and interviewed about the crime by the police, the fear that the truth about their innocence may never come out, and perhaps negative reactions from family members, colleagues, neighbours, etc. Under such circumstances, relevant questions may well elicit strong physiological responses.

Consider the case of Roger Keith Coleman who was convicted of the rape and brutal murder of his sister-in-law and sentenced to death (Ford, 1995). Coleman insisted that he was innocent, and there were some weaknesses in the case against him. For example, after his conviction, four persons came forward to testify that they heard someone else confess to the crime. In a last attempt to prove his innocence, Coleman asked for a polygraph examination. The polygraph test was administered 12 hours before the scheduled execution. It was announced that Coleman failed the test, and he was put to death later that night. It is perhaps not surprising that Coleman failed the polygraph test. It is difficult to believe how Coleman could have been anything other than extremely aroused when answering the crucial relevant questions during this polygraph test, whether he was guilty or innocent.

Ekman (1985/2001) cites four more reasons why innocent suspects may do the opposite of what is expected in CQT theory, and show stronger responses to the relevant than to the probable lie questions. First, innocent suspects may think that the police are fallible. Innocent suspects who have been asked to take a polygraph test know that the police have already made one mistake by suspecting them of a crime that they did not commit. Perhaps they have already tried to convince the police of their innocence, without any success. Although, on the one hand, they could see the test as an opportunity to prove their innocence, on the other hand it is also feasible that they could fear that those who made the mistake of suspecting them will make further mistakes. In other words, if police methods are fallible enough to make them falsely appear suspicious, could their polygraph tests not also be fallible? Second, the innocent suspect may think that machines are fallible. For example, they may have had difficulties with their own personal computer or other technical equipment, and therefore do not believe that a machine can ever be infallible. Third, the innocent suspect may think that the police are unfair. People may dislike or distrust the police and will therefore expect and fear that the polygraph examiner will misjudge them or cheat them. Finally, the innocent suspect may be a fearful person. Someone who is generally fearful might respond more to the relevant questions than to the probable lie questions.

Not only is it unclear why innocent suspects should necessarily show larger responses to relevant questions than to probable lie questions, it is also unclear why guilty examinees should necessarily respond more strongly to relevant questions than to probable lie questions. First, they could successfully control their physiological responses during the test so that they do not show a "guilty pattern", as I will discuss in further detail in the "countermeasures" section below. Second, according to CQT logic, the more serious the transgressions, the stronger the physiological response the examinee will show when lying about them. Thus, a guilty examinee who has committed serious crimes in the past, is lying about serious transgressions in the probable lie questions, which will thus evoke strong physiological responses during these questions, and the physiological responses during the relevant questions may not exceed these responses. Third, if the efforts of the examiner during the pretest phase of the CQT are successful, then guilty examinees may also be concerned about the probable lie questions. Why should a guilty suspect necessarily be less concerned about the probable lie questions than about the relevant questions, given that the suspect is under the impression that a deceptive response to those probable lie questions may be harmful to his or her case? (Ben-Shakhar, 2002; Ben-Shakhar & Furedy, 1990).

Proponents of the CQT acknowledge the theoretical problems of their test. They report that "...there is no specific-lie response or pattern of reactions that is peculiar to deception" (Raskin & Honts, 2002, pp. 4–5), but also that "...the assessment of the underlying causal nature of the physiological responses is a separate scientific question from determining the accuracy of a test. It is quite possible, and acceptable, to have a test validated as accurate for its specified purpose. . . . without having a complete theoretical understanding of the underlying theoretical constructs" (Raskin & Honts, 2002, p. 2). I think this point of view is problematic as I discuss later on in this chapter.

Lack of incorporation of psychological knowledge in the CQT protocol

Several scholars report that CQT polygraphers fail to incorporate in their protocol psychological knowledge gained over the years (Fiedler, Schmid, & Stahl, 2002; Mitchell, 2002; National Research Council, 2003). It is believed that this is the result of only a few scientists being actively engaged in research on physiological lie detection (Iacono, 2000; National Research Council, 2003). With regard to knowledge that is disregarded, Fiedler *et al.* (2002) refer to literature that suggests that preceding questions codetermine the meaning or interpretation of subsequent questions. In alignment with this, Cullen and Bradley (2004) found that accurate detection of deception depends on the position of the probable lie questions in the sequence. In addition, Fiedler *et al.* (2002) argue how vaguely phrased probable lie questions may be interpreted differently by different examinees. Moreover, Mitchell (2002) refers to persuasion literature that could possibly be used to convince examinees about (i) the effectiveness of the polygraph in detecting deception and (ii) the importance of the probable lie questions.

Perhaps the most severe criticism of the theoretical development of polygraph testing comes from the National Research Council (2003). They noted that, although psychophysiological lie detection is one of the oldest branches of applied psychology, it has been affected relatively little by advances made over more than a century in basic psychology and physiology.[8] For example, The National Research Council noticed that some polygraphers still record skin resistance rather than skin conductance, whereas it has been known for decades that the latter gives a more useful measure of EDA. The National Research Council also questioned the multi-measurement approach used by polygraph examiners. In their chart coding, they do not distinguish between the different physiological measurements (EDA, blood pressure, and respiration) but simply average the scores of the different measurements

[8]This conclusion seems at odds with the opinion of the then UK Home Secretary, David Blunkett, who approved a scheme in the United Kingdom where sex offenders face polygraph tests before and after being freed from prison. He said of this scheme: "We are all a bit sceptical because we've all been brought up with the spy films and the way in which the KGB are allegedly able to train people to avoid them [polygraph]", but, he continued "We are talking about really modern technology in the 21st century and we are testing it" (*The Independent*, 29 May 2004, p. 4).

to obtain a total score. Yet, it has been known for a long time that electrodermal (EDA), cardiovascular (blood pressure), and respiratory measures respond in different ways to various psychological states (see also Kemeny, 2003).

The National Research Council also refers to literature about "stigmas" that has been ignored by polygraphers but that could predict which innocent individuals run a heightened risk of being falsely accused. Stigmatised individuals are members of socially devalued groups. Stigmas can be caused by gender, skin colour, bodily deformity, sexual orientation, etc. Stigmatised individuals often feel threatened during interactions with nonstigmatised individuals, and typically display large physiological responses during task performance but not during baseline or resting periods (National Research Council, 2003).

A crucial and difficult role for the examiner: lack of standardisation in conducting the test

Preparing a polygraph test might be considered a very sophisticated piece of psychological engineering (Lykken, 1998). For the polygraph exam to work, the examiner should formulate probable lie questions that, on the one hand, should elicit in innocent suspects *stronger* physiological responses than the relevant questions. On the other hand, in guilty suspects, these probable lie questions should elicit *weaker* physiological responses than the relevant questions. Obviously, it is not easy for the examiner to formulate questions that meet these criteria. If examiners make an examinee very concerned about the probable lie questions they will run the risk of deception not being detected in guilty examinees. In such a case, physiological responses to probable lie questions may be equal to physiological responses to relevant questions, resulting in an inconclusive test outcome. On the other hand, if examiners do not embarrass examinees enough with the probable lie questions, they run the risk of assessing innocent examinees as deceptive, because in such a case their physiological responses to the relevant questions may be higher than their physiological responses to the probable lie questions.

Another problem with formulating probable lie questions is that the examinee's responses to these questions are not known lies but "probable" lies (Lykken, 1998). The examiner believes that the examinee's answers to these questions are untrue, but can often not be absolutely certain of this. Obviously, when the assumptions made by the examiner are incorrect the probable lie questions will not achieve the desired effect, since in that case the examinee is actually telling the truth when answering these questions.

Raskin acknowledges the difficulty in formulating probable lie questions. In his own words: "The traditional CQT is difficult to administer, and the level of psychological sensitivity, sophistication, and skill of the examiner are crucial to obtaining an accurate outcome. Unfortunately, many polygraph examiners lack adequate training in psychological methods and do not understand the basic concepts and requirements of a standardised psychological test. These problems are exacerbated when the examiner formulates and introduces the control[9] questions to the subject, because it is very difficult to standardise the wording and discussion of the questions for all subjects. A great deal depends on how the subject perceives and responds to the control questions when they are introduced and discussed during the pre-test interview" (Raskin *et al.*, 1989, p. 8). In other words, it all depends on the quality of the polygraph examiner, and Raskin does not appear to be impressed by the average quality of them. Barland (1988) has also expressed his concerns about the quality of many polygraphers in the United States. It is this problem perhaps that Iacono (2000) refers to when he stated that one of the reasons why the CQT is controversial is that the training of polygraph examiners occurs outside the scientific environment.

The important role of the examiner may lead to a situation where the examiner's belief about the examinee's guilt prior to the polygraph test will influence the test outcome. The examinee is not a complete stranger to the examiner, who will know important details about the examinee (including case-file information) and will form an impression of him or her in the pre-test interview where the probable lie and relevant questions are formulated. If the examiner believes that the suspect is *innocent*, he or she may put the suspect under much pressure during the probable lie questions, increasing the likelihood that the examinee will pass the test. In contrast, if the examiner believes that the suspect is *guilty*, he or she may not put so much emphasis on the probable lie questions, increasing the chance of the examinee failing the test. The result will be that the test outcome reflects the examiner's prior beliefs about the examinee's guilt.

This may well have happened in a CBS case study. In 1986, CBS (a US television station) enlisted the help of four polygraph-testing firms in New York to ascertain which of four of its employees had stolen a valuable camera. As the polygraph examiners arrived (on different days) to do the testing, each one was told that the manager's suspicions were focused on one particular individual; a different individual was specified for each polygraph examiner.

[9] Probable lie questions were called control questions in the past.

This was a charade. There was no missing camera and the four employees were aware of this. They were merely instructed to deny that they had stolen any camera (i.e., to tell the truth). As an incentive they were each offered US $50 if they succeeded in passing the polygraph test. The outcome was that each of the four polygraphers positively and confidently identified a culprit, and in each case that was the one employee whom that polygrapher had been told was suspected as being the culprit (Lykken, 1988).

Apart from beliefs about the possible guilt of examinees, other impressions about examinees may also influence polygraphers, and hence polygraph outcomes, such as liking the examinees or feeling sorry for them. In these circumstances "guilty outcomes" are perhaps less likely to occur. Examiners may produce predisposed polygraph outcomes deliberately (Honts & Perry, 1992), but not necessarily so. Social psychology research has convincingly demonstrated that people can be unintentionally affected by pre-information or their prior beliefs (Rosenthal, 1976; Rosenthal & Rubin, 1978). The examiner's expectations about the outcomes of the test can change their behaviour slightly in terms of speech pauses, tone of voice, voice loudness, etc. when asking the relevant questions during the test. Examinees could pick up these signs, consciously or unconsciously, resulting in a heightened physiological reaction to these questions (Ben-Shakhar, 2002).

The important message of this section is that CQT examinations cannot be seen or presented as an objective and scientific process. Instead, the outcomes of the test very much depend on the skills, biases, and expectations of the individual examiner.[10]

A crucial and difficult role for the examiner: lack of standardisation in scoring the charts

In the global charts scoring approach, the examiner forms a global impression of the examinee's guilt on the basis of the polygraph charts and combines this impression in some unspecified manner with non-polygraphic sources of information such as case-file information, the examinee's criminal history, or the examinee's demeanour throughout

[10] The notion that the examiner can influence CQT outcomes plays a role in the "friendly polygrapher" debate. Sometimes suspects ask and pay for their own polygraph test to prove their innocence. CQT opponents argue that examiners have a natural inclination to serve the interest of their clients, which makes it more likely that the suspect will pass the test (Iacono, 2000; Lykken, 1998). CQT proponents claim that their own files show no evidence for a friendly polygrapher effect (Raskin & Honts, 2000). However, the subjective nature of the CQT means that the friendly polygrapher hypothesis cannot be rejected on theoretical grounds.

the test, to reach an overall decision. This means that the decision-making process is subjective, as it depends on the examiner. Moreover, it is not verifiable. Other people cannot determine why a particular examiner has come to a particular conclusion.

Backster and Raskin acknowledged this problem and introduced the numerical scoring approach. However, even this approach is subjective because it is not standardised, as I illustrated in Figure 11.2. A lack of standardisation means that two examiners who score the same chart could reach different decisions, and this actually happens. Carroll (1988) reported that inter-agreement rates between different examiners in laboratory studies ranged from 0.61 to 0.95. These scores are satisfactory but show that differences between examiners do occur.[11] Moreover, as Fiedler *et al.* (2002) indicated, differences between examiners do not occur randomly but are most likely to happen in cases where the outcomes are not clear-cut.

Research has demonstrated that in cases that are not clear-cut, contamination can occur between an examiner's expectations about the examinee's guilt and scoring of the chart via the numerical scoring approach. Experienced polygraph examiners were requested to interpret polygraph charts of a certain examinee (Elaad, Ginton, & Ben-Shakhar, 1994). Prior to this chart-coding task, information about the examinee was given. Some examiners were told that the examinee ultimately confessed to being responsible for the crime, while others were informed that another person had confessed to the crime. These manipulated prior expectations affected the examiners' decisions about the guilt of the examinees, and these prior expectations were likely to be supported. However, this happened only if the polygraph charts did not include clear indications of guilt or innocence. This type of contamination also takes place in real-life settings, as I will discuss later.

The issue of subjectivity in coding polygraph charts emphasises, again, the important role of the examiner. These problems could be overcome by using a computer-based method for scoring polygraph data, and such methods have been developed (Kircher & Raskin, 1988, 2002; Olsen, Harris, Capps, & Ansley 1997). Decisions made via computer programs are objective, but whether they are accurate depends on the computer algorithms that are used. The National Research Council (2003)

[11] These scores refer to agreement between examiners in interpreting the charts. However, this is not the only agreement aspect that matters in polygraph testing. As I explained above, formulating the questions is not standardised in the CQT. This means that two examiners could formulate different questions. The issue thus is whether two examiners who set up two different CQTs for the same examinee would reach the same verdict. This issue has not been addressed in the polygraph literature (Iacono, 2000; Ben-Shakhar, 2002), but the CBS study presented in the main text suggests that it would be appropriate to do so.

reported that little is known about the accuracy of such algorithms (but see Kircher & Raskin, 2002, for a description).

Whatever is done to make the chart coding more objective, it is unlikely to please all CQT opponents. One of those critics, John Furedy, expressed his opinion about computerised scoring of polygraph charts in the following way: "Data based on flawed procedures will still yield flawed output . . . or in less technical and polite words: garbage in, garbage out" (Furedy, 1996a, p. 57).

Vulnerability and illegality of using deceptive procedures

Misleading the examinee plays a crucial role in the CQT. First, examinees are led to believe that strong responses to the probable lie questions will lead the examiner to decide that the examinee was *deceptive* with respect to answering the relevant questions about the crime under investigation, whereas, in fact, the opposite is true: in such circumstances the examinee will be judged as *truthful*. Second, examinees are led to believe that the test is infallible which, as we will see below, is untrue. Third, the stimulation test, used to demonstrate the polygraph's infallibility, is misleading. In the stimulation test, a card test is used, but this is not a concern-based polygraph test but a polygraph test based on the orienting reflex, described in Chapter 12. In other words, a fundamentally different test (a test based on the orienting reflex) is used to demonstrate the infallibility of the concern-based CQT.

One may argue that it is appropriate to use deceptive procedures. Supporters will probably say that the end justifies the means, and that it is important to convict serious criminals, if necessary by deceiving them. Proponents may also argue that polygraph tests sometimes benefit innocent suspects, namely when the test confirms their innocence.

Opponents may point out that deceiving suspects is always inappropriate and can have negative consequences. For example, it may undermine public confidence and social trust in the police and other agencies that conduct polygraph tests; or suspects may think that they are allowed to lie since the polygraph examiner is allowed to lie to them. Also, suspects may decide to stop cooperating with the investigators when they find out that they have been fooled. However, cooperation is often needed to obtain further evidence because polygraph outcomes typically do not count as evidence in court. In a similar vein, someone who realises that he has been cheated during an investigation may opt for non-cooperation in future investigations.

Perhaps more importantly, the deceptive procedures make the CQT vulnerable because the examinee must believe these lies to make the test effective. Thus examinees must believe that lying to the probable lie questions will look suspicious, and that the test is infallible. According to Elaad (1993) and Lykken (1988) it is unlikely that all examinees will believe this. There are probably dozens of books and articles written about the CQT, including details about the nature of probable lie questions, and the fallibility of the test. Even popular newspaper articles appear about the CQT (Furedy, 1996b). Of course, those who are taking polygraph tests have access to this literature (especially with access to the internet which is now so ubiquitous) and may well have read (some of) it. If they know about the procedures and/or fallibility of the test, they will not believe the examiner's lies about the importance of the probable lie questions and the infallibility of the test.

Finally, it is important to consider whether it is *legal* to use deceptive procedures. The aspect of legality does not play a role in polygraph discussions, because it is dominated by American scholars. In the US the police are allowed to lie to suspects, but the situation is different in other countries, including many West European countries, where carrying out deceptive investigative procedures is not acceptable under law. Conducting a CQT may thus be illegal in some countries.

Countermeasures

As I explained earlier in this book, the moment people realise that someone is going to judge whether or not they appear convincing, they may well try to influence their responses in such a way that they actually do make a credible impression. Therefore, examinees sometimes attempt to influence polygraph outcomes and try to produce physiological responses which will lead the examiner to conclude that they are not deceptive. Such attempts are called countermeasures. Although in principle both truth tellers and liars could attempt to employ countermeasures, liars are most likely to do so, as they have the tendency to take their own credibility less for granted (Chapter 3). If countermeasures could be shown to be effective, it may have major implications for polygraph testing, because it will make the test less accurate.

Different countermeasures can be distinguished: physical countermeasures include tongue biting or foot tensing (by pressing the toes against the floor), and mental countermeasures include thinking about a frightening experience or counting backward. These countermeasures will result in physiological responses that will be detected by the polygraph. By employing countermeasures during the probable lie questions, examinees can artificially increase their physiological reactions to these questions, thereby increasing their likelihood of passing the test.

Reid and Inbau (1977) are not concerned about the effectiveness of countermeasures. They argue that it is highly improbable that countermeasures can succeed because a properly trained examiner would notice that the examinee is trying to fool them. However, as CQT critics and proponents acknowledge, research has shown that countermeasures can be successfully applied (Ben-Shakar & Dolev, 1996; Elaad, 1987, cited in Ben-Shakhar & Furedy, 1990; Honts & Alloway, 2007; Honts & Amato, 2002; Honts, Devitt, Winbush, & Kircher, 1996; Honts, Hodes, & Raskin, 1985; Honts, Raskin, & Kircher, 1987, 1994). In one experiment, participants were trained for 30 minutes in the use of either physical or mental countermeasures (Honts *et al.*, 1994). After this training session they underwent a CQT. The examiner, who was experienced in polygraph testing, only detected 12% of those employing physical countermeasures, whereas none of the mental countermeasure users raised the suspicion of the examiner. The physical and mental countermeasures were equally effective, each enabling approximately 50% of the participants to beat the polygraph test. In another study, 70% of guilty participants were classified as innocent following training to press the toes and bite the tongue, whereas the test results of the remaining 30% of trained guilty participants were inconclusive (Honts *et al.*, 1987). In other words, none of the guilty participants who used countermeasures were classified as guilty. In contrast, 80% of the guilty participants who did not use countermeasures were classified as guilty. These findings contradict Reid and Inbau's claim that experienced examiners will detect the use of countermeasures.

CQT critics argue that countermeasures are relatively easy to learn (Ben-Shakhar, 2002; Lykken, 1998). A countermeasures test that supports this view was conducted by Floyd "Buzz" Fay, a man who was falsely convicted of murder on the basis of a failed polygraph examination. He took it on himself to become a polygraph expert during his two and half years of wrongful imprisonment. He coached 27 inmates, who all freely confessed to him that they were guilty, how to beat the CQT. After only 20 minutes of instruction, 23 of the 27 inmates were successful in defeating the polygraph examination (Ford, 1995; Kleinmuntz & Szucko, 1984).

The Directed Lie Test

I explained that one of the major criticisms of the CQT is the lack of standardisation in conducting the test. Therefore, much depends on the skills of the examiners, but, as CQT proponents themselves acknowledge, many polygraph examiners lack the necessary skills (Honts, 2004; Raskin & Honts, 2002). In an attempt to standardise the CQT, the Directed Lie Test (DLT) has been developed (Horowitz, Kircher, Honts, & Raskin, 1997; Raskin & Honts, 2002; Raskin, Honts, & Kircher, 1997). The DLT is identical to the CQT with one exception. In a DLT, the probable lie questions are substituted by so-called directed lie questions that are standardised and can be asked in all situations. Typical examples of directed lie questions are "In the first 27 years of your life, have you ever told even one lie?" and "Before age 27, have you ever broken even one rule or regulation?" (Raskin & Honts, 2002). The examiner introduces these directed lie questions by informing examinees that they are designed to establish how the examinee physiologically responds when he or she lies. The examinee is further told that the examiner will compare their physiological responses to the directed lie questions with their physiological responses when asked about the crime under investigation. Examinees will be instructed to answer "No" to the directed lie questions and to think about particular situations in which they did tell a lie or did break a rule.

The rationale behind the DLT is similar to that behind the CQT. Guilty suspects are thought to be mostly concerned with the relevant questions and are expected to show the strongest responses to these questions. Innocent suspects are thought to be most concerned about the directed lie questions. They will focus on demonstrating that their physiological reactions when they lie (i.e., to directed lie questions) indeed differ from their physiological reactions when telling the truth (i.e., to relevant questions).

Criticism Concerning the Directed Lie Test

Undoubtedly, the DLT is a more standardised test than the CQT. However, the DLT procedure does not address any of the other criticisms related to the CQT and is thus equally vulnerable to these criticisms: a weak theoretical foundation; lack of incorporating psychological knowledge in the test procedure; lack of standardisation in scoring the charts; vulnerability and illegality of using deceptive procedures;[12] and vulnerability to countermeasures. In some aspects, the DLT may be worse than the CQT. Ben-Shakhar (2002) pointed out that the DLT has been disputed even among CQT supporters, for example, because the excessive emphasis placed on the directed lies could lead to liars remaining undetected. Ben-Shakhar (2002) further argues that the directed lies are clearly distinct from the relevant questions and therefore cannot serve

[12] In the DLT no lies are told about the nature of the directed lie questions, however, lies are still told about the infallibility of the test, and a test based on the orienting reflex is still used to demonstrate the efficacy of the concern-based DLT.

as controls for intrapersonal differences. Iacono and Lykken (1999) argued that explaining to examinees the purpose of the directed lie questions, and clarifying the importance of giving strong responses to these questions, may guide those who attempt to beat the test.

ACCURACY OF CQT POLYGRAPH EXAMINATIONS

Case studies about the accuracy of polygraph testing are easy to find in the polygraph literature. However, they do not give us real insight into the accuracy of polygraph testing. To obtain real insight, laboratory and field studies need to be carried out. Virtually all of these studies are related to two types of tests: the Guilty Knowledge Test (Chapter 12) and the CQT. In this section I will discuss the accuracy of the CQT.

To test the accuracy of the CQT both laboratory studies and field studies have been conducted. In previous chapters I have already discussed the differences between laboratory and field studies. In most CQT laboratory studies, liars committed a mock crime and were instructed to deny this wrongdoing when asked about it in the subsequent CQT. Truth tellers did not commit the mock crime and therefore could honestly deny wrongdoing when asked about it during the CQT. The advantage of laboratory studies is that researchers know who is lying and who is telling the truth. A further advantage is that researchers could examine the impact of several factors on the accuracy of a CQT, such as the impact of employing countermeasures.

A disadvantage of laboratory studies is that the consequences of passing or failing the test are not as high as they typically are in real life. Participants may be promised a reward if they can convince the examiner of their innocence, or they may be threatened with punishment if they cannot. Sometimes these punishment threats seem serious. For example, Bradley and Janisse (1981) threatened their guilty and innocent participants with a "painful but not permanently damaging electric shock" if they failed the test. However, the stakes in laboratory studies are almost inevitably lower than in most real-life examinations. For this reason, some scholars have argued that laboratory studies are of no use in estimating the accuracy of the CQT in applied settings (Ben-Shakhar, 2002; Elaad & Ben-Shakhar, 1997; Kleinmutz & Szucko, 1982; Lykken, 1988, 1998). Others have suggested that laboratory studies may be useful as long as certain criteria are met, such as using representative participant populations (not only college students), and carrying out realistic polygraph practices (including using expert examiners and introducing some motivation for the examinee to deceive the examiner) (Kircher, Horowitz, & Raskin, 1988; Raskin, 1989).

Those who argue against laboratory based CQT polygraph examinations have a valid point to make: the possibility that innocent examinees are more concerned when answering the relevant questions than when answering the probable lie questions grows as the seriousness of the event referred to in the relevant questions increases, and the more serious the allegation, the greater the extent of concern. Laboratory studies cannot deal with serious allegations. As such, they mask the important problem of innocent examinees showing larger responses to the relevant questions than to the probable lie questions and therefore can be seen as optimum testing conditions for conducting a CQT.

Field studies deal with CQT examinations of suspects in real-life criminal cases. The advantage of field studies is that they are realistic. The major disadvantage is the ambiguity about ground truth, that is, establishing with certainty the actual guilt or innocence of the examinees. Ideally, this would be achieved via corroborative and conclusive evidence that is gathered independently from the polygraph test, such as DNA evidence. However, this type of evidence is typically not available in cases where polygraph tests are carried out, because if there had been strong proof of guilt or innocence in an actual case, polygraph tests are unnecessary (Honts, Kircher, & Raskin, 2002). Therefore, confessions are widely accepted as a method of establishing the ground truth in field studies.

Using confessions as ground truth is problematic due to the *sampling bias* (Fiedler, Schmid, & Stahl, 2002; Patrick & Iacono, 1991). The sampling bias can (i) explain why polygraph examiners typically believe that they are highly accurate and (ii) make it possible that the accuracy scores obtained in CQT field studies are inflated. Let's start with the perceptions of the polygraph examiners. After an examinee fails a polygraph test, the polygraph examiner and police will believe that the examinee has committed the crime and will attempt to extract a confession. Some suspects will confess and these cases will all bolster the view of the examiner that the suspect was actually guilty. However, the confession could have been false,[13] but that is difficult to discover for the examiner or police. It will only be discovered if evidence subsequently turns up that discredits the false confession. Since the police

[13] Innocent suspects often falsely confess when they see little further opportunity to convince the jury or judges of their innocence and decide to confess in order to receive a lesser sentence (Gudjonsson, 2003). Alternatively, it might be that as a result of failing a polygraph test suspects actually believe that they are guilty. This happened to Peter Reilly and Tom Sawyer, and their cases are described in Box 11.2. Suspects may start to doubt their innocence because they trust the polygraph and believe that it is infallible. This may sound naive, but bear in mind that such infallibility is assured by the examiner. Moreover, the police will tell the suspect afterwards that the outcome was accurate. Some suspects believe this.

believe the confessor to be guilty, they will not be searching for this additional evidence and so it is unlikely to arise. Not all suspects who fail the test will confess, and confessions may be particularly unlikely from those who are innocent. Since there is often no other evidence available in the case, these cases will remain unsolved. However, unsolved cases do not discredit the belief that the guilty polygraph verdict was actually correct, it just means that the case could not be proven. The result is that it is unlikely that an examiner will be confronted with evidence that discredits his or her "guilty" polygraph verdict.

What about suspects who pass the test? Because the examiner and police believe that these suspects are innocent, they are unlikely to be submitted to a grilling interrogation, and may even not be interviewed at all. Moreover, because the suspect passed the test, there is no evidence against him or her. Suspects who pass the test, including guilty suspects, are thus unlikely to confess. Since the police believe that someone other than the suspect who just passed the test has committed the crime, they may pursue their search for evidence and for a new suspect. The search is unlikely to be successful in cases where the suspect who passed the polygraph test was actually guilty. Many of these cases therefore remain unresolved, and unresolved cases do not discredit the innocent polygraph verdict. The result is that it is unlikely that an examiner will be confronted with evidence that discredits his or her "not guilty" polygraph verdict.

The same reasoning explains why accuracy rates in CQT field studies could easily become inflated. In a field study only cases are included where the examinee has confessed (these examinees will be considered guilty) or where someone other than the examinee has confessed (these examinees will be considered innocent). Guilty suspects who fail the test (correct polygraph decisions) are more likely to confess than guilty suspects who pass the test (incorrect polygraph decisions) and therefore more likely to be included in the field study. Guilty suspects who failed the test are most likely to confess for the reasons given in the previous paragraphs: the police will put more effort in eliciting confessions from guilty suspects who failed the test because they believe that these suspects are guilty; and guilty suspects who failed the test are confronted with evidence against them (the failed polygraph test), whereas guilty suspects who passed the test do not have evidence against them.

Let's give an example with numbers of how inflation may occur in guilty examinees (derived from the National Research Council, 2003, p. 115). Suppose that of 300 guilty suspects 200 fail the polygraph test and 100 pass it. This means that two-thirds of guilty suspects are correctly classified which equals an accuracy rate of 67%. Further suppose that the police manage to elicit a confession from 30% of those who failed the test and from 10% of those who passed the test. It means that 66 guilty suspects who failed the test confessed (30% of 200 suspects) and that 10 guilty suspects who passed the test confessed (10% of 100 suspects). Only those 76 guilty suspects will be included in the field study because they are the only ones to have confessed. The accuracy rate in this subsample is 87% (66 correct decisions out of a total of 76 decisions), which is far higher than the actual accuracy rate of 67%.

For innocent examinees, the chance that someone else will confess (requirement for the innocent examinees to be included in the field study sample) increases when the innocent examinee passes (correct polygraph decision) rather than fails (incorrect polygraph decision) the test. In the latter case, the police will think that the innocent examinee is the culprit and will put little effort in finding another suspect.[14]

Box 11.2 Discredited polygraph verdicts contributing to false confessions: the Peter Reilly and Tom Sawyer cases

Peter Reilly. Eighteen-year-old Peter Reilly returned home one night to find his mother dead. He realised that she had been murdered and he called the police. The police interviewed Reilly and suspected him of killing his mother. They conducted a polygraph test. The police told him that he failed the test, thus indicating that he was guilty even though he had no memory of the incident. Transcripts of the interrogation sessions revealed that Reilly underwent a remarkable transformation from denial to confusion, then to conversion ("Well, it really looks like I did it") and finally to a written confession. Two years later, independent evidence revealed that Reilly could not have committed the murder, and that the confession that even he had come to believe was false (Kassin, 1997).

Tom Sawyer. Tom Sawyer's next-door neighbour was murdered by manual strangulation. The police became suspicious of Sawyer solely because he seemed to be nervous when they spoke to him during a routine interview that they held with him as one of the murdered woman's neighbours. Sawyer was invited to the police station for a second interview. In response to questions about his general background, Mr Sawyer discussed both his social anxiety and the fact that he had been an alcoholic. In trying to engage Mr Sawyer in conversation about the crime, the detectives asked him to help them to create a scenario of how the murder might have

[14] See Pollina, Dollins, Senter, Krapohl, and Ryan (2004) for a study where laboratory and field data were compared.

happened. Mr Sawyer, who loved to watch detective shows on television, was eager to help and joined in. The police let Mr Sawyer explain several scenarios and accused him at the end of having committed the murder. The police claimed that Mr Sawyer mentioned nine facts that only the killer could have known. Subsequent analysis of the interrogation transcripts showed that all of the crucial information was introduced into the interrogation by the police.[15]

Following the accusation, Mr Sawyer strongly denied his guilt. The police obtained fingerprint and hair samples from him and suggested a polygraph examination. Mr Sawyer believed that the polygraph examination would prove his innocence and agreed with the examination. After the test, the examiner told Mr Sawyer that the test proved he was lying. Mr Sawyer's confidence began to erode. He was no longer able to express firm denials of guilt, and could only say that he still did not believe he had committed the crime. His main defence against the police's demands to confess was his lack of memory of having committed the crime. The police replied by saying that he was blocking out his memory of the crime and that he had had a blackout, just as he had often experienced when he had been drinking. At this point, Mr Sawyer still refused to accept fully that he had committed the crime, and pinned his hopes on the other tests (fingerprints and hair test) revealing his innocence. The detectives decided to lie to him and told him that his hair samples matched hairs found on the victim's body.

On receipt of this information, Mr Sawyer's resistance collapsed. He agreed that "all the evidence was in" and that he must have committed the crime. During the next period of the interrogation the detectives wanted to obtain an accurate description of the crime, which was impossible for Mr Sawyer because he had not committed the crime. The police helped Mr Sawyer somewhat by suggesting to him what they knew had happened or what they thought might have happened. For example, the police believed that the victim had been sexually assaulted. Encouraged by the police detectives, Mr Sawyer confessed to having raped the victim. However, when the medical examiner's report was received, no evidence of sexual assault was indicated (Ofshe, 1989).

[15] Being confused about the source of information occurs frequently and is called a *source monitoring error* (Raye & Johnson, 1980). People sometimes become confused about whether they themselves or others presented some information, or about which out a number of people presented the information. This sometimes happens during meetings. People sometimes disagree afterwards about who said what during a meeting.

CQT Accuracy Rates: Laboratory Studies

In principle, six outcomes are possible in a CQT. A *guilty suspect* may show stronger physiological responses to relevant questions than to probable lie questions, in which case he or she has failed the test, which is an accurate outcome. It is also possible that they show similar responses to relevant and probable lie questions, in which case the outcome is inconclusive. Finally, they may show stronger physiological responses to probable lie questions than to relevant questions. In this case they passed the test, which is an incorrect outcome.

An *innocent suspect* may show stronger physiological responses to probable lie questions than to relevant questions, in which case the suspect has passed the test, which is the correct outcome. It is also possible that the innocent suspect shows similar responses to probable lie questions and relevant questions, in which case the outcome is inconclusive. Finally, the innocent suspect may show stronger responses to relevant questions than to probable lie questions, and hence has failed the test, which is the incorrect outcome.

In short, there are two types of error. First, the classification of a guilty examinee as innocent, which is called a *false-negative error*. Second, the classification of an innocent examinee as guilty, which is called a *false-positive error*. Both types of error are undesirable, and which is the most serious depends on the circumstances. In criminal investigations (investigations examining theft, rape, murder, etc.) a false-positive error is probably the most serious error, as this may result in an innocent suspect being prosecuted and convicted, whereas a false negative error means that the guilty suspect will escape prosecution and conviction. In Western legal systems often the principle "It is better to acquit ten guilty persons than to convict a single innocent" is used (de Keijser & van Koppen, 2007). Polygraph tests are also carried out in security threat investigations to catch spies (see below). In those circumstances a false-negative error may be the most serious type of error because it will result in a spy being undetected. Obviously, one undetected spy can cause a lot of harm.

Table 11.2 provides an overview of the six reviews of CQT laboratory studies that have been published in English that I am aware of. Each review author included only those laboratory studies that they felt met minimum quality criteria[16] rather than all those available to them.

[16] The National Research Council (2003) felt that the quality of polygraph research is often low. They examined 194 separate studies (a mixture of laboratory and field studies) of which only 57 (50 laboratory and 7 field studies) met their minimal standards of scientific adequacy. The National Research Council's analyses are not included in the tables in this chapter because no accuracy rates were reported. The National Research Council (2003, p. 4) described the accuracy of polygraph tests "...well above chance, though well below perfection".

Table 11.2 Results of laboratory study reviews examining the accuracy of CQT polygraph examinations

	Guilty condition			Innocent condition		
	Guilty	Innocent	Inconclusive	Guilty	Innocent	Inconclusive
Ben-Shakhar & Furedy (1990, $n = 9$)	80%	7%	13%	15%	63%	22%
Honts (1995, $n = 8$)	77%	10%	13%	8%	84%	8%
Honts (2004, $n = 11$)	82%	7%	11%	10%	83%	7%
Kircher *et al.* (1988, $n = 14$)	74%	8%	18%	12%	66%	22%
OTA report (Saxe *et al.*, 1985, $n = 12$)	74%	7%	19%	16%	60%	24%
Raskin & Honts (2002, $n = 11$)	80%	8%	12%	8%	84%	8%

Note: n = number of studies reviewed

This is why different reviews included different numbers of studies and yielded somewhat different outcomes.

The results regarding CQT laboratory examinations show that these tests are rather accurate in classifying guilty examinees. Depending on the review, a majority of guilty examinees (74% to 82%) were correctly classified and relatively small numbers of guilty examinees (7% to 10%) were incorrectly classified as innocent. A less positive picture emerges for innocent examinees, at least in some reviews. In three reviews only 60% to 66% of innocent examinees were correctly classified and between 12% and 16% of innocent examinees were falsely accused of lying. However, in the reviews conducted by Honts and Raskin, two proponents of the CQT, the results are more positive with 83–84% correct classifications of innocent suspects and 8–10% false accusations of innocent suspects.

CQT Accuracy Rates: Field Studies

The seven reviews of CQT field studies that have been published in English that I am aware of are presented in Table 11.3. They involve accuracy rates obtained by examiners who did not carry out the actual examinations but only had access to the polygraph charts (e.g., independent evaluators). The different reviews produced somewhat different outcomes. Different authors included different studies in their reviews, leaving out field studies that they felt did not meet minimum quality criteria. Saxe, Dougherty, and Cross (1985) attempted to provide "an objective description, to the extent that is possible, of current psychological knowledge of polygraph testing" (p. 356). They presented a review that was initiated by the US congressional Office of Technology Assessment (OTA) to advise President Reagan about polygraph testing. They found 10 studies that met the OTA standards, and their review included more studies than any other review published to date. Confessions were used as ground truth in most studies included in the reviews, as was the case with the studies included in the OTA review. Given the uncertainties about the ground truth, the reported accuracy rates can only be considered to be estimates. Moreover, as I explained above, the accuracy rates of field studies could easily be inflated due to the sampling bias.

There is overlap between the reviews regarding the correct classification of guilty examinees. Depending on the review, between 83% and 89% of the guilty suspects were correctly classified. Differences emerged between reviews regarding the number of incorrect decisions made regarding guilty suspects (i.e., classifying a guilty suspect as innocent). Depending on the literature review, error rates vary from 1%

Table 11.3 Results of field study reviews examining the accuracy of CQT polygraph examinations

	Guilty condition			Innocent condition		
Authors	Guilty	Innocent	Inconclusive	Guilty	Innocent	Inconclusive
Ben-Shakhar & Furedy (1990, $n = 9$)	84%	13%	3%	23%	72%	5%
Carroll (1999, $n = 3$)[1]	83%	17%	–	47%	53%	–
Honts & Perry (1992, $n = 3$)	86%	11%	3%	30%	59%	11%
Iacono & Patrick (1997, $n = 3$)[2]	84%	–	–	–	56%	–
Lykken (1998, $n = 4$)[3]	86%	–	–	–	61%	–
OTA report (Saxe *et al.*, 1985, $n = 10$)	87%	11%	2%	19%	75%	6%
Raskin & Honts (2002, $n = 4$)[4]	89%	1%	10%	12%	59%	29%

Note: n = number of studies reviewed
[1]Inconclusive cases were not included; [2]Incorrect classifications and inconclusive cases were not reported; [3]Incorrect classifications and inconclusive cases were not reported; [4]Honts (2004) is identical to Raskin and Honts (2002) and therefore not reported

to 17%. The lowest error rate was reported by Raskin and Honts,[17] and the highest by Carroll, a CQT opponent.

There is even less agreement amongst the reviews regarding innocent suspects. However, in every review the accuracy rates for innocent suspects are lower than the accuracy rates for guilty suspects. Depending on the review, between 53% and 75% of innocent suspects were correctly classified and between 12% and 47% were incorrectly classified. Again, the lowest error rate was reported by Raskin and Honts, and the highest error rate by Carroll.

The pattern that emerges in CQT field studies in particular is that the CQT is vulnerable to false-positive errors, that is, the false accusation of innocent suspects. This is perhaps not surprising given that such errors occur when an examinee shows stronger physiological responses to relevant questions than to probable lie questions. Apparently, despite the examiner's efforts, innocent suspects do not always show more concern for the probable lie questions than for the relevant questions. The percentages of false-positive errors are higher in field studies than in laboratory studies, which is not surprising either. In field studies the stakes are high, and in these tests in particular suspects will realise the major consequences that the relevant questions have for them.

As mentioned above, the results presented in Table 11.3 were derived from evaluations of the physiological data by examiners who did not conduct the polygraph tests themselves (e.g., independent evaluators). Raskin and Honts (2002, p. 32/33) noticed that although "this is a desirable practice from a scientific viewpoint. . .it is usually the original examiner who makes the decision on how to proceed in an actual case". They then argue that to obtain a real picture of the accuracy of a CQT the verdicts of original examiners should be examined rather than those of independent evaluators. Their review revealed that original examiners are typically more accurate than independent evaluators (Raskin & Honts, 2002). For example, Honts, Raskin, Kircher, and Hodes (1988) obtained very high accuracy rates with original examiners: their accuracy rates for guilty (92%) and innocent (91%) classifications are amongst the highest accuracy rates ever reported in deception research.

There are two possible reasons as to why original examiners are more accurate than independent evaluators. First, original examiners could be more skilful in scoring charts. This explanation is unlikely, as the independent evaluators used in polygraph field research are usually

[17] Raskin and Hont's (2002) review includes Patrick and Iacono's (1991) field study. Iacono (2000) claims that the accuracy scores from his study are misrepresented in that review.

experienced and skilful polygraphers. Second, original examiners use extra non-polygraph information to make their decision, such as the examinee's demeanour during the test and case-file information, and this additional information increases accuracy. CQT proponents acknowledge that this happens (Honts, 1996; Horvath, 1973; Raskin & Honts, 2002). In his study, Honts (1996) found that in four cases – involving two innocent and two guilty suspects – the examiner made a decision on the basis of the polygraph outcomes, whereas a decision should not have been made because the test outcomes for these four cases were inconclusive. In all four cases, however, the decisions made by the examiners (guilty or innocent) were correct, indicating that extra non-polygraph information led the examiners to the correct decision.

In their field study, Patrick and Iacono (1991) examined in more detail the extent to which decisions made by polygraph examiners were entirely based on chart data or on a combination of chart data and non-polygraph information (see also Iacono, 2000). Available to the researchers were (i) the original examiners' numerical chart scores, and (ii) the original examiner's final verdicts. In theory, the original examiners' numerical chart scores should dictate the final verdicts but this was not the case. For the 132 tests that were numerically scored as deceptive, only 82% were diagnosed as deceptive and 18% received a truthful verdict. Another 69 tests were numerically scored as inconclusive, but 45% of these suspects received a truthful verdict and 4% were diagnosed as deceptive. These findings suggest that original examiners take into account information other than chart scoring. Also, there was a pattern in the decisions they made. In 93% of the cases where the examiners' verdict was not in alignment with their numerical scoring, a truthful verdict was given.

The researchers had another source of information. Also available to them were numerical chart scores given by independent evaluators. Comparisons between numerical scoring of original examiners and independent evaluators revealed that even in their original scoring examiners were biased towards truthfulness of the examinees. The numerical scores of 72 suspects received a truthful verdict amongst original examiners whereas the numerical scores of only 51 suspects received a truthful verdict amongst independent evaluators. Importantly, original examiners were more often correct in their final verdicts (90% accuracy) than in their numerical scoring (70% accuracy). The independent evaluators obtained the lowest accuracy (55%).

In summary, those who had only access to the polygraph charts (a real test of the accuracy of the CQT) achieved only 55% accuracy, considerably lower than the original examiners who had access to all available information (90% accuracy). It also appeared that the original examiners' verdicts and numerical scores were adjusted to favour truthful verdicts. This point is interesting because, as we just saw, the CQT is vulnerable to making false-positive errors. Examiners appear to correct their verdicts somewhat to avoid this error.

Box 11.3 A unique polygraph field study

A unique attempt to conduct a polygraph study in a realistic setting and maintaining certainty about the ground truth was made by Ginton, Daie, Elaad, and Ben-Shakhar (1982). The participants in this study were 21 Israeli policemen who took a paper-and-pencil test that was presented as a requirement for a police course in which they were enrolled. They were asked to score their own tests, which provided an opportunity to cheat by revising their initial answers. However, the test answer sheets were chemically treated so that cheating could be detected. It turned out that seven out of the 21 participants cheated.

Later, all were told that they were suspected of cheating. They were offered a polygraph examination, and were told that their future careers in the police force might depend on the outcome of this examination. (The option to allow the police officers to refuse to take the test was realistic. As mentioned before, in criminal investigations taking a polygraph test is an option and not a requirement for a suspect.) Although initially all 21 policemen agreed to undergo a polygraph examination, one guilty officer did not turn up for the actual examination, and two (one guilty and one innocent) refused to take the polygraph test. Three other guilty participants confessed just before the polygraph interrogation,[18] so the final sample included only two guilty and 13 innocent participants. The CQT was used, and the outcomes were rather accurate. Both guilty officers were accurately detected, but two of the 13 innocent officers were mistakenly judged to be lying.

Although the findings appear to give impressive support for the CQT, in fact they do not. The examiner who made the decisions about guilt and innocent of the examinees had access to both polygraph data and the participants' general demeanour during the examination. His accuracy with regard to innocent participants

[18] Ekman (1993) pointed out that these figures support what polygraph examiners claim, namely that the threat of taking a polygraph exam does produce confessions among guilty suspects. However, the findings also suggest that refusal to take the test is no certain guarantee of guilt.

TSA 15-00014 - 002002

(11 out of 13 correct) was equal to the accuracy of an observer who watched the polygraph examinations but who had no access to the polygraph data. A blind evaluator who had access only to the polygraph data classified only seven out of 13 cases correctly, and classified three innocent examinees as guilty and three outcomes as inconclusive (Carroll, 1991).

PERSONALITY AND THE POLYGRAPH

People may think that it is difficult to detect lies told by psychopaths with a polygraph. Psychopaths have low EDA responses (Lorber, 2004) and this may make it difficult to distinguish between probable lie and relevant questions. In addition, psychopaths are less bothered than normal people about danger in general and punishment in particular (Lykken, 1998; Porter & Woodworth, 2007). It might therefore be that psychopaths will not be concerned during a polygraph test, making it impossible to detect their lies. However, research suggests that psychopaths and non-psychopaths are equally detectable with a polygraph (Patrick & Iacono, 1989; Raskin & Hare, 1978).

Other studies examined differences between the detectability of introverts and extraverts with polygraph testing and found no differences either (Steller, Haenert, & Eiselt, 1987; Watson & Sinka, 1993). Also studies investigating background characteristics such as gender and race did not reveal noticeable differences (National Research Council, 2003). However, the National Research Council (2003) reported that it would be too premature to conclude that personality traits and background characteristics have no effect, as not enough good quality studies have been carried out examining this.

(PRE)EMPLOYMENT SCREENING

Specific Incident, Non-Specific Incident, and Pre-Employment Tests: Differences in Ambiguity

In the US polygraph tests are conducted in the area of employment and pre-employment screening. Although the Employee Polygraph Protection Act of 1988 substantially restricted the use of the polygraph in the private sector, it is still used in the public sector, such as by the US government, the FBI, and various police departments (Krapohl, 2002).

Screening of employees is carried out for different purposes. First, to obtain insight into a specific security-threatening incident. Suppose there is evidence that a computer that contains high security information has been tampered with, and that only a limited number of employees have access to that computer. Those employees can be asked to undergo a polygraph test to examine whether they are responsible for tampering with the computer. In such a situation a *specific incident* CQT can be carried out, similar to the tests I described earlier in this chapter. For example, a relevant question in this case could be: "Did you tamper with the computer?"

Second, polygraph tests are also used to examine whether employees have committed or are engaged in sabotage, or have provided classified information to unauthorised persons, and so on, without having any evidence that such illegal acts have actually taken place (National Research Council, 2003). The difference with the computer tampering example above is that there is no specific incident the examiner can refer to. The examiner has therefore no option other than to phrase relevant questions that are vague and *non-specific*. A relevant question that could be asked in a *non-specific incident* CQT is "During the last five months, have you ever provided classified information to an unauthorised person?" Asking such a non-specific question creates problems due to its ambiguity. The question may elicit memory about a specific incident in one employee, but not in another employee who has experienced a similar incident. Also, suppose that the employee fails the test. What does that reveal? What is the employee accused of? The polygraph test does not provide an answer because it did not ask a specific question. The employee needs to be asked in a post-test interview why a heightened physical response occurred, and whether his or her answer during the post-interview is truthful would remain to be seen. For example, an employee who had done something illegal may then recall a minor violation that he or she knows will have no further consequences to cover up for a more serious violation.

Polygraph tests are also conducted for screening in application procedures. In this situation the test is used to predict particular aspects of future job performance, such as the chances that the applicant, if employed, will commit security violations. These *pre-employment tests* are even more problematic than non-specific incident tests because they are even more ambiguous than the non-specific incident tests. In this case, questions about the (applicant's) past are used to predict the future. Sometimes questions about the past (e.g., "Have you ever used drugs in your life?") do not cover the same topic as the future behaviour an employer may be interested in (e.g., Will this person be a future spy?).

The National Research Council (2003), whose task it was to evaluate the use of polygraph tests in (pre)employment settings, noticed that hardly any research, and no quality field research, has been carried out regarding the accuracy of non-specific incident and pre-employment tests. However, they concluded that due to the ambiguity of non-specific incident and pre-employment tests, accuracy in those types of tests are almost certainly lower than what can be achieved in specific incident polygraph tests.[19]

Base Rates

Apart from ambiguity, the *base rates* issue is another problem with pre-employment and employment screening. Base rates refer to the occurrence of deception in the sample to be tested. For example, in laboratory studies, half of the participants are lying, thus the base rate of lying is 50%. However, in pre-employment settings the base rate is considerably lower. For example, it may well be that only 10 out of 1000 applicants are potential spies or terrorists. The problem with low base rates is that the lie detection test needs to be highly accurate to have any diagnostic value. Suppose that a polygraph test classifies 80% of the liars and 80% of the truth tellers correctly. If all 1000 applicants were tested, eight (potential) spies would be correctly identified, but also 198 honest applicants (20% of 990 honest applicants) would be classified as deceptive and thus considered to be potential spies.

Krapohl (2002) does not seem to worry much about these false accusations. He argues that the applicant pool typically exceeds the number of vacancies. Hence, if there are 500 jobs available and the employer is guided by the polygraph test, there are in the example still 794 applicants to choose from, and this sample excludes the eight potential spies. This reasoning should raise ethical concerns, as 198 applicants will not be hired, not because they have done anything wrong or are not equipped for the job, but because the polygraph examiner made a mistake. Also, as I reported earlier, these 198 unfortunate individuals may include a large proportion of stigmatised people, because they typically display large physiological responses during task performance but not during baseline or resting periods (National Research Council, 2003).

Let's now consider testing current employees. The base rate of employees who spy on their employer could be equally low as the base rate of potential spies who apply for jobs. It may well be that only 10 out of 1000 employees are spying on their employer. If the polygraph test classifies truths and lies with 80% accuracy, eight spies will be correctly identified, but also 198 honest employees will be suspected of spying,

[19]See Ben-Shakhar (2002) for problems in formulating comparison questions in non-specific incident and pre-employment tests.

and the polygraph charts cannot distinguish between the spying and falsely accused employees.

One can imagine morale levels in a workplace where 20% of employees are suspected of spying. Therefore, polygraph users may wish to avoid making false accusations, for example by making the threshold higher for being classified as a liar. This will certainly reduce the number of employees that are falsely accused of spying but will also increase the likelihood that employees who are spying will not be caught. The National Research Council (2003, p. 6) therefore concluded that "Polygraph testing yields an unacceptable choice. . .between too many loyal employees falsely judged deceptive and too many major security threats left undetected. Its accuracy in distinguishing actual or potential security violators from innocent test takers is insufficient to justify reliance on its use in employee security screening in federal agencies".

Other Objectives: The Polygraph Mystique

Conducting polygraph tests with applicants and employees may be useful for other objectives. As reported above, people often believe that a polygraph test is infallible; the National Research Council (2003) labels this the "lie detection mystique". Polygraph tests may therefore work as a deterrent, discouraging potential spies from applying for a job and current employees from spying if they know that they will be polygraphed. Polygraph testing may also result in spontaneous admissions of wrongdoing. Employees may be more honest in non-polygraph interviews if they know that they will be later subjected to a polygraph test. The ability of polygraph tests to induce confessions is considered as one of the reasons why polygraph tests are carried out and accepted in various settings (Lykken, 1998; Cross & Saxe, 2001).[20]

Due to their image of being infallible, polygraph tests may also increase public confidence in national security organisations. However, since polygraph tests are in fact fallible the same confidence in polygraph testing may present a danger to national security services if such organisations rely too heavily upon them (National Research Council, 2003). Lykken (1998) believes that this is the case. He gives the example of Aldrich Ames, the CIA agent who sold secrets to the Soviets for many years and who passed several polygraph tests during this time.

[20] To achieve confessions, carrying out a proper polygraph test is not necessary. Participants who were hooked up to an impressive looking electronic device, introduced as a lie detector but in fact a "pile of electronic junk" became more truthful in their subsequent interviews. (The so-called bogus pipeline effect, Cross & Saxe, 2001; Jones & Sigall, 1971.)

He stated that Ames "... .succeeded in his spy career for as long as he did because his ability to beat the lie detector deflected official suspicions" (Lykken, 1998, p. 3).

POLYGRAPH TESTS WITH SEX OFFENDERS

The number of polygraph tests conducted with sex offenders is rapidly increasing. They are on the rise in the United States (Branaman & Gallagher, 2005; Consigli, 2002; Honts, 2004), and have been introduced, for example, in the Netherlands (Meijer, 2004) and the United Kingdom (Wilcox, Sosnowski, & Middleton, 2000). It is perhaps not surprising that attention is paid to sex offenders in particular. Sex offenders typically minimise the nature and extent of their activities and interests; they are likely to re-offend; and their offences are considered serious (British Psychological Society, 2004). Amongst other applications, polygraph screening of sex offenders is used as a condition of their release of prison (Honts, 2004) or to monitor their behaviour once released (Kokish, Levenson, & Blasingame, 2005). Such monitoring tests are sometimes carried out as regularly as every three to six months (Cross & Saxe, 2001).

Different types of polygraph tests are carried out with sex offenders (Branaman & Gallagher, 2005). In a *specific incident* polygraph test, known offenders are polygraphed regarding their involvement in a specific allegation (e.g., "Did you touch John's penis"?). Those tests are similar to the crime-related specific incident CQT examinations discussed above in this chapter. More frequently used is the *maintenance* polygraph test (.e.g., "Have you violated any conditions of your probation?" or "Have you done anything over the last three months that would concern your probation officer"?) and *sexual history disclosure* polygraph tests (e.g., "Have you withheld any information from your sexual history?"). These are non-specific incident polygraph tests and I already discussed the problems of such tests. The questions are (unavoidably) vague which makes them ambiguous. Such tests are therefore almost certainly less accurate than specific incident tests.

There is no research available regarding the accuracy of the maintenance and sexual history disclosure tests. The problem is obtaining ground truth (Grubin, Madsen, Parsons, Sosnowski, & Warberg, 2004). Some sex offenders pass the test suggesting that they have not been involved in deviant sexual behaviour or high-risk behaviour during the period covered by the polygraph examination. However, ground truth regarding these cases is lacking and it could well be that these offenders have been engaged in deviant acts despite passing the test. Other offenders fail the test and may admit acts of deviant sexual behaviour or types of high-risk behaviour in a subsequent interview, but it is unknown whether the acts they admit to did actually happen. Perhaps they admit to minor acts to cover up more serious acts. Alternatively, they may falsely confess to deviant sexual acts or high-risk behaviour, as self-reports of examinees have revealed (Kokish *et al.*, 2005).

Despite the lack of knowledge about their accuracy and the doubts raised about their accuracy, the sex offenders polygraph tests are popular amongst practitioners. One reason for this popularity is that they lead to admissions of deviant sexual acts and high-risk behaviour that were previously unknown to their probation officers (Ahlmeyer, Heil, McKee, & English, 2000; Consigli, 2002; Grubin *et al.*, 2004). It is further claimed that these polygraph tests are "the missing link in preventing recidivism" (Consigli, 2002, p. 239). I already discussed above the confession-inducing effect and possible deterrent effect of the polygraph. However, as I just mentioned, we cannot be certain whether the admissions are truthful, neither do we know how many deviant sexual acts and instances of risk behaviour remain uncovered by polygraph testing.

Some More Concerns

A few more concerns about polygraph testing of sex offenders are worth mentioning. First, in the monitoring programmes, repeated testing of the same offender is the norm (Consigli, 2002; Cross, & Saxe, 2001). The impact of multiple testing on the accuracy of polygraph tests is unknown, but it may well have a damaging effect. It could lead to habituation effects over time, and could also increase the use of countermeasures. Second, inevitably, if polygraph tests are conducted on a voluntary basis some sex offenders will choose not to take part. How do you judge those who refuse to participate? We know that refusing to participate does not necessarily imply guilt (Ginton *et al.*, 1982), but probation officers are likely to interpret a refusal as a sign of guilt. For example, Don Grubin, who is in charge of the sex offender polygraph trial in the United Kingdom, reported that "The guys who have taken the tests are the ones who want to prove that they are no risk. A lot of guys we want to test, the ones we are concerned about, will not do it" (*The Times*, 29 May 2004, p. 7). Third, the knowledge of supervision officers about polygraph testing is a concern. Branaman and Gallagher (2005) reported that supervision officers erroneously presume that specific incident and other types of polygraph tests are equally accurate.

Table 11.4 Answers to the five *Daubert* questions for CQT polygraph assessments

	CQT laboratory	CQT field
(1) Is the scientific hypothesis testable?	Yes	Problematic
(2) Has the proposition been tested?	Yes	Possibly
(3) Is there a known error rate?	Yes	Yes, too high
(4) Has the hypothesis and/or technique been subjected to peer review and publication?	Yes	Yes
(5) Is the theory upon which the hypothesis and/or technique is based generally accepted in the appropriate scientific community?	No	No

LEGAL IMPLICATIONS

In this section I will discuss what are, in my view, the implications of this chapter's findings for the use of CQT outcomes as scientific evidence in criminal courts. As I explained in Chapter 8, I will use the set of guidelines provided by the United States Supreme Court for admitting expert scientific evidence in the federal (American) courts (*Daubert v. Merrel Dow Pharmaceuticals, Inc.*, 1993). Table 11.4 summarises my verdict.

Question 1: Is the Scientific Hypothesis Testable?

The underlying rationale of the CQT is that liars will show stronger physiological responses to relevant questions than to probable lie questions and that truth tellers will show the opposite pattern and show stronger physiological responses to probable lie questions than to relevant questions. This premise can easily be tested in laboratory studies, but the findings may lack ecological validity. Field studies may therefore provide a more appropriate test of the CQT rationale, but are difficult to conduct due to difficulty in obtaining the ground truth. My answer to the first *Daubert* question is thus "yes" regarding CQT laboratory tests and "problematic" regarding CQT field tests.[21]

[21] My verdict about this question could easily have been more negative, due to the fact that the CQT is not based on a valid theoretical premise. The question which then arises is, as phrased by Gallai (1999, p. 96): "How can the underlying science be tested when it has been conceded that no adequate explanation exists as to how the underlying science operates?"

Question 2: Has the Proposition been Tested?

A substantial number of CQT laboratory studies have been carried out, thus my answer to the second question is "yes" regarding CQT laboratory studies. Numerous field studies have been published to date, but they are subject to debate. The problem is that the quality of these published polygraph field studies is low (National Research Council, 2003), and most problems are related to establishing a ground truth that meets scientific standards. Regarding CQT field tests, my answer to the second *Daubert* question is therefore "possibly".

Question 3: Is there a Known Error Rate?

The error rate in CQT laboratory studies is known and, depending on the literature review, varies between 7% and 10% for false negatives (incorrect classification of guilty suspects) and between 8% and 16% for false positives (incorrect classification of innocent suspects). However, due to the lack of ecological validity of such studies, it may be better to examine the error rates in field studies. The exact error rate in field studies is unknown, because of the lack of ground truth and the presence of a sampling bias in most of these studies. Different literature reviews provide different estimates of the error rates. They vary from 1% to 17% for false negatives and 12% to 47% for false positives. Given these sometimes high error rates, particularly for innocent suspects, allied to the knowledge that these error rates could easily be deflated due to the sampling bias, my verdict is that these error rates are too high to present CQT outcomes as being "beyond reasonable doubt".

Another argument that prevents me from drawing the conclusion that CQT outcomes are beyond reasonable doubt is that in a CQT much depends on the skills of the individual examiner. There are likely to be large individual differences in the accuracy scores of different CQT examiners. It is therefore inappropriate for individual CQT examiners to use the accuracy rates of field studies as an indication of their own accuracy rates. A given individual examiner may obtain accuracy rates that are well below these known accuracy rates.

Question 4: Has the Hypothesis and/or Technique been Subjected to Peer Review and Publication?

My answer to the fourth *Daubert* question is "yes" regarding both CQT laboratory and field research, although the number of high-quality CQT field studies is relatively low.

Question 5: Is the Theory upon which the Hypothesis and/or Technique is Based Generally Accepted in the Appropriate Scientific Community?

Iacono and Lykken (1997) published a survey where scientific opinion concerning the polygraph was examined. They asked members of the American Society of Psychophysiological Research (who can be considered as experts) and Fellows of the American Psychological Association (Division 1, General Psychology) for their opinions regarding the CQT. The opinions of both groups of psychologists were very similar. A minority of interviewees (approximately 33%) considered the CQT to be based on scientifically sound psychological principles or theory. Moreover, when asked whether they would advocate that courts admit CQT outcomes as evidence, only approximately 22% agreed for when the test outcome indicates guilt and only approximately 25% agreed for when the test outcome indicates innocence. My answer regarding the final *Daubert* questions is thus "no".[22]

Overall Verdict

The CQT does not meet the *Daubert* guidelines for admitting expert scientific evidence in criminal courts. There are problems with testing its underlying rationale; doubts about the quality of the field studies that have been carried out, particularly regarding the ground truth obtained in such studies; the error rates are probably too high, particularly for innocent suspects; and the test is disputed in the appropriate scientific community. Gallai (1999), Honts (1994), and Saxe and Ben-Shakhar (1999) also carried out a *Daubert* analysis regarding CQT polygraph examinations, similar to my review in this section. My verdict complies with Gallai's (1999) and Saxe and Ben-Shakhar's (1999) verdicts but contradicts Honts' (1994) verdict who believes that the CQT meets the *Daubert* criteria.

VOICE STRESS ANALYSIS AND THERMAL IMAGING

Voice Stress Analysis

The basic assumption of voice stress analysis is that liars experience more psychological stress than truth tellers (Gamer, Rill, Vossel, &

[22]This survey was carried out by two CQT critics (Iacono and Lykken), and the results are not generally accepted by CQT proponents (Honts, 2004). Honts (2004) refers to three other surveys that all yielded more positive results regarding the CQT. However, the Iacono and Lykken survey is the only survey to have been published in a peer reviewed quality journal (*Journal of Applied Psychology*), although that does not guarantee in itself that the results are reliable and valid.

Gödert, 2006). This psychological stress results in minor changes in the blood circulation, which subsequently influences various characteristics of the voice. Therefore, specialised equipment called Voice Stress Analysers (VSA) (also sometimes called Psychological Stress Evaluators) use microphones attached to computers to detect and display voice indices such as intensity, frequency, pitch, harmonics, or micro tremors (see http://www.polygraph.org/voicestress.htm for useful information about VSA). Sometimes voice stress analysis is introduced as an alternative technique to polygraph testing but this is misleading because the only difference lies in the physiological responses which are measured (voice characteristics versus EDA, blood pressure, and respiration). The VSA microphones measure physiological responses non-intrusively and this has obvious benefits. Data can be gathered more quickly (the examinee does not have to be connected to a machine) and covertly, that is, without the examinees being aware that they are being assessed.

However, voice stress analyses have severe limitations. First, they may detect truths and lies inaccurately. For example, Horvath (1978, 1979) and Gamer *et al.* (2006) could not distinguish between truth tellers and liars on the basis of voice stress analyses, whereas the same truth tellers and liars were detected above chance levels when EDA measures were taken. Second, the CQT, the only type of test that is unanimously supported by proponents of concern-based lie detection, cannot be carried out without the examinee's awareness, because (i) the questions need to be discussed with the examinees prior to the examination, and (ii) background information about the examinee is necessary to formulate probable lie questions (see above). This implies that, if voice stress analyses were to be carried out covertly (e.g. by insurance companies when they discuss claims with clients over the telephone), a CQT cannot be carried out.

An RIT could be used, but this test is disputed even by supporters of concern-based lie detection due to its lack of controlling for intrapersonal differences.[23] Alternatively, just single questions could be asked (i.e., "Are you positive that your sunglasses have been stolen"?) but this is even less reliable than an RIT, because no control questions are asked. As a result, there is even no attempt to control for intrapersonal differences, and, uniquely for concern-based lie detection, there is no attempt to control for interpersonal differences either. Unsurprisingly, the National Research Council (2003, p. 167) concluded that "although proponents of voice stress analysis claim high levels of accuracy,

[23] See Zraick, Gentry, Smith-Olinde, & Gregg (2006) for the relevance of controlling for intrapersonal differences when measuring pitch of voice.

empirical research on the validity of the technique has been far from encouraging".[24]

Thermal Imaging

Thermal imaging is a technique whereby changes in temperature patterns (and thus blood flow) around the eye are detected via special cameras. The assumption behind the technique is that liars will show instantaneous warming around the eyes as part of a fright/flight response. One thermal imaging lie detection study in particular has attracted extensive media attention to date (Pavlidis, Eberhardt, & Levine, 2002a). I think that there are two reasons for this attention. First, the study was published in the journal *Nature*, a highly respected journal that often attracts wide media attention. Second, the authors claimed in their article that the technique has "an accuracy comparable to that of polygraph examination. . .and has potential for application in remote and rapid screening, without the need for skilled staff or physical contact" and "it could be used for instantaneous lie detection without the subject even being aware of the test" (Pavlidis *et al.*, 2002a, p. 35).

Airport screening comes to mind when reading these claims. If these claims were correct then the thermal patterns around the eyes of every passenger could be non-intrusively measured and the terrorists amongst them could be instantly detected. However, for such a test to work the thermal patterns should be the equivalent of Pinocchio's growing nose, because only then would all potential terrorists fail and all honest passengers pass the test. Unfortunately, thermal imaging is not the equivalent of Pinocchio's growing nose, and the test, as suggested by Pavlidis and colleagues, is doomed to fail. No attempts are made to control for inter- and intrapersonal differences and that makes the test inaccurate. Perhaps some potential terrorists may be aroused and fail the thermal imaging test. However, not all of them will be, and those who are not will be allowed to board the plane. The likelihood of a potential terrorist passing the test becomes higher if terrorist organisations find out that thermal imaging procedures are used. They can simply recruit individuals who they know are likely to pass thermal-imaging screening tests.

In contrast, individuals who have no intention of doing any harm but who are naturally nervous will fail the test. Further, individuals who are excited may fail the test. They could be excited about the act of

[24] Despite this negative VSA verdict, John Hutton, the Work and Pensions Secretary in Great Britain, is introducing voice analyses to tackle fraudulent persons (*The Independent*, 6 April 2007).

flying; about going on holiday; about just spotting an attractive person, and so on. Also, airports are probably places full of individuals who are anxious for all manner of reasons, and individuals who are stressed may fail the test. They could be stressed because they are afraid of flying; are worried about missing the plane; are on their way to an important but problematic business meeting, and so on.

In summary, thermal imaging tests are likely to target vast numbers of innocent passengers who all have to undergo further scrutiny to show that they have no intention of blowing up an aircraft. Unsurprisingly, the National Research Council (2003, p. 157) concluded that thermal imaging "does not provide acceptable evidence for the uses of facial thermography in the detection of deception".[25,26]

EVALUATION OF CONCERN-BASED POLYGRAPH TESTS

The CQT is the only concern-based polygraph test that has the support of all proponents of concern-based polygraph testing. The rationale of CQT is that liars will be more concerned about relevant questions than about probable lie questions and therefore will show the strongest physiological responses to the relevant questions. In contrast, truth tellers will be mostly concerned about the probable lie questions and therefore will show the strongest physiological responses to these probable lie questions. Problems with the CQT include its weak theoretical foundation; lack of standardisation in conducting the test and scoring the charts; and its vulnerability to countermeasures.

The accuracy of CQT in real life is not really known due to the lack of ground truth in the CQT field studies that have been published. Moreover, the accuracy rates reported in these field studies could easily be inflated due to the sampling bias. However, the accuracy rates of CQT laboratory studies are typically high and it therefore sounds reasonable to conclude that CQT can discriminate between truths and lies above the level of chance.

CQT outcomes are generally not presented as evidence in criminal courts and I think it should stay this way. In my view, the CQT does not meet the *Daubert* criteria and two problems are particularly noticeable.

[25] Pavlidis, Eberhardt, and Levine (2002b, p. 602) published an erratum stating that some problems associated with their technique "*might* preclude large-scale application". In my view, this erratum does not go far enough.

[26] More recently the thermal imaging technique has been used to detect deception in a mock crime laboratory study using the CQT protocol. The benefit of thermal imaging compared to a CQT polygraph examination is that arousal can be measured non-intrusively. The outcomes were comparable to laboratory CQT polygraph studies (Tsiamyrtzis, Dowdall, Shastri, Pavlidis, Frank, & Ekman, 2007).

First, the test is vulnerable to false-positive errors, that is, the false accusation of innocent suspects. Introducing CQT verdicts in criminal courts could thus result in a substantial number of innocent suspects being falsely convicted. Second, the CQT has a weak theoretical foundation. CQT proponents seem to acknowledge the theoretical problems with their test, but believe that it is acceptable to use a test without having a complete understanding of the underlying theory (Raskin & Honts, 2002). I think this point of view is problematic. The benefit of a sound theory is that someone can predict when a test does and does not work. Take for example the X-ray equipment used in airport security screening. Since there is full understanding about how materials are represented in X-ray images, one can accurately predict how different types of material show up on the X-ray screens. Without a theory, no predictions can be made. Perhaps certain people are at high risk of failing the CQT, thereby increasing the likelihood that they will be falsely accused, and perhaps others are at low risk of failing the CQT, thereby increasing the likelihood that they will be undetected. Also, perhaps the CQT is more suitable to use for some types of crime than for other types of crime. Without a theory we do not know.

Apart from not having met the *Daubert* criteria, the second reason why CQT outcomes should not be admissible in criminal courts is that it is not standardised (Ben-Shakhar, 2002; Ben-Shakhar, Bar-Hillel, & Lieblich, 1986). The CQT is not an objective test, and, for example, CQT outcomes can be influenced by the beliefs examiners have about the guilt of examinees prior to the test. These initial beliefs could be based on non-factual evidence that typically would not be accepted as evidence in criminal courts, such as the criminal history of the examinee, hearsay, and as on. However, if CQT outcomes did become admissible, such non-factual evidence would be introduced in criminal courts by means of the polygraph verdict.

Not only should the CQT be inadmissible in criminal courts, I would not encourage criminal investigators to use the CQT either. I have three reasons for this. First, proponents may argue that it is "just another tool in the box" for the investigator, but I disagree. From discussions with CQT examiners it became clear to me that such examiners strongly believe that the technique is highly accurate, and as I explained in this chapter, this could be an exaggeration caused by the sampling bias. However, the result is that CQT users probably perceive the CQT as *the* tool rather than *a* tool in the box, and this could result in relying too much on its outcomes. Second, its vulnerability to false-positive errors makes the use of CQT in criminal investigations unwarranted. Although a false accusation by the police or polygraph examiner may not result in a wrongful conviction, it still could have serious consequences for the suspect, such as being interviewed more often, having to spend more time in police custody, not being able to convince family members, friends, and colleagues of being innocent, etc. Third, due to the lack of standardisation of the CQT, much depends on the skills of individual examiners. Even CQT proponents acknowledge that many examiners lack appropriate skills. Introducing the CQT thus runs the risk of many tests being improperly or poorly conducted due to the incompetence of examiners.

The CQT is also used in employment and pre-employment screening and, increasingly, with sex offenders. Using the CQT in such settings is even more problematic than in criminal investigations. The main problem is that there is often no specific incident about which the examiner can formulate relevant questions for the test. Relevant questions are therefore ambiguous, which makes the test outcomes less accurate. Proponents argue that polygraph examinations in these situations may elicit admissions or work as a deterrent. However, one cannot judge the reliability of these admissions as long as the ground truth in these cases is unknown. Moreover, one wonders for how long a test with obvious limitations can work as a deterrent?

CHAPTER 12

Physiological Lie Detection: The Orienting Reflex Approach

This chapter discusses a polygraph test called the Guilty Knowledge Test (GKT), also known as the Concealed Information Test. The GKT was developed and described by David Lykken (1959, 1960, 1988, 1991, 1998). In this chapter I explain the rationale behind the GKT, how the test works, and the main problems associated with it. This chapter reveals that the GKT has a sound theoretical foundation and is popular amongst scientists. The chapter further reveals that the GKT cannot be used in all situations, which is often given as the reason why practitioners in most countries seldom use it (Bashore & Rapp, 1993). However, Israel and, in particular, Japan are exceptions. The GKT is the most frequently used polygraph test in Japan, and there are about 70 polygraphers who carry out a total of 5000 GKT polygraph examinations each year, all in criminal investigations. The polygraph outcomes are accepted as evidence in Japanese criminal courts (Nakayama, 2002).

I also review in this chapter research examining the GKT's accuracy in classifying truth tellers and liars, and discuss whether GKT outcomes should be accepted as evidence in criminal courts. It will become clear that GKT laboratory studies have revealed high accuracy rates in classifying truth tellers and liars, but that GKT field studies have showed considerably lower accuracy rates, particularly in pinpointing liars. The low lie accuracy rates obtained in field studies is one of the reasons why I think that GKT verdicts should not be used as evidence in criminal courts. However, I do believe that this test can be used as a lie detection tool in criminal investigations.

THE GUILTY KNOWLEDGE TEST

The GKT is based on the orienting reflex (Pavlov, 1927; Sokolov, 1963). An orienting response (I will use the words reflex and response interchangeably) occurs when someone is confronted with a personally significant stimulus. An often used illustration is the cocktail party phenomenon (Cherry, 1953): people can be unaware of conversations going on around them, yet notice when their name is mentioned in one of these conversations. An orienting reflex probably occurs to facilitate an adaptive response to the stimulus by the individual (National Research Council, 2003; Sokolov, 1963). Orienting reflexes result in physiological responses measured by the polygraph, such as an increase in electrodermal activity (EDA) (Nakayama, 2002), and a decline in heart rate (Verschuere, Crombez, de Clercq, & Koster, 2005), and also result in the occurrence of P300 brain waves measured via EEGs (Rosenfeld, 2002). In most GKT research and daily practice a polygraph is used, I will return to the EEG-P300 research later on in this chapter.

The orienting theory can be applied to lie detection. Suppose that someone was killed with a knife which was left at the crime scene. As long as information about the crime has not been leaked out, only the murderer could know which type of knife was found at the crime scene. In a typical GKT examination, examinees are attached to the polygraph and will be shown several knives, including that used in the murder. For each knife the examinee will be asked whether he or she recognises it as the one used in the crime. The examinee is instructed to answer "No" each time this question is asked. The murderer is familiar with the murder weapon and is therefore likely to show orienting reflexes when the knife used in the crime is shown; innocent examinees have no knowledge about which knife was used and will therefore show no orienting reflexes.[1]

Ben-Shakhar, Bar-Hillel, and Kremnitzer (2002) discuss how a GKT could have been used in the John Demjanyuk trial. Demjanyuk was convicted for Nazi war crimes in Israel in 1988 because, it was claimed, he was Ivan the Terrible from the concentration camp Treblinka. However, his defence team argued that this was a case of mistaken identity

[1] Strictly speaking, the GKT does not claim to detect lies, but restricts itself to measuring orienting reflexes. Therefore, several researchers refer to this test as a recognition test rather than a lie detection test.

Table 12.1 An example of the Guilty Knowledge Test

(1) You know that Nicole has been found murdered, Mr. Simpson. How was she killed?	
–	Was she drowned?
–	Was she hit on the head with something?
–	Was she shot?
–	Was she beaten to death?
–	Was she stabbed?
–	Was she strangled?
(2) Where did we find her body? Was it:	
–	In the living room?
–	In the driveway?
–	By the side gate?
–	In the kitchen?
–	In the bedroom?
–	By the pool?

Source: taken from Lykken (1998, p. 298)

and that the defendant was not Ivan the Terrible. Material uncovered after the verdict raised doubts about the conviction, and the Israeli Supreme Court reversed the guilty verdict. Perhaps a GKT could have established whether or not Demjanyuk was Ivan the Terrible. There are biographical details in each person's life that are probably remembered by that person but seldom known to strangers, such as the maiden name of one's mother. A GKT where Demjanyuk had been given several names including the maiden name of Ivan the Terrible's mother could have shed light on whether Demjanyuk was Ivan the Terrible.

A GKT could perhaps also be used to identify some Al Qaeda members (National Research Council, 2003). For example, those members who have visited the training camps in Afghanistan may still remember information about the location of the camps and striking features inside the camps. Those aspects may be known to GKT examiners but probably not known to many other people. Examiners could thus formulate questions about training camp characteristics.

Lykken (1998) described how a GKT could have been conducted in the O.J. Simpson murder case, directly after the body of Simpson's wife was found. The GKT is presented in Table 12.1.[2]

Examinees often give different physiological responses to the first alternative than to the other alternatives in a GKT. The first alternative is therefore usually a filler (i.e., not the correct alternative) (Bashore & Rapp, 1993). Theoretically, examinees may show orienting responses to

[2] The card test, often used in CQT as a stimulation test (Chapter 11), is another example of a GKT.

the correct alternative just by coincidence. The chance that this may have happened in one of the questions in the O.J Simpson example is one out of five (disregarding the first alternative). However, this chance goes down to one out of 25 that this happens to questions 1 and 2. Hence, the more questions asked, the lower the chance of an examinee showing orienting responses to all the correct alternatives just by coincidence. It is therefore important to ask multiple questions in a GKT. In Japan typically seven questions are asked, each with five alternatives (four fillers and one correct alternative) (Nakayama, 2002). The chance that an examinee will display orienting reflexes by coincidence to the correct alternatives in all seven questions is one in 16,384 when the first alternative is always disregarded.

It is further possible that when the examiner knows the correct alternatives to the questions this will, consciously or not, influence his or her behaviour each time the correct alternative comes up during the test. For example, an examiner, keen to know whether the examinee is the culprit, may become slightly excited each time he or she mentions the correct answer. A change in the examiner's behaviour may be noticed, consciously or not, by the examinee, which may influence the examinee's physiological responses (Rosenthal & Rubin, 1978). It is therefore preferable that the examiner who conducts the actual test does not know which the correct alternatives are.

CRITICISM CONCERNING THE GUILTY KNOWLEDGE TEST

Several problems have been outlined with using the GKT. In this section I will discuss the most important ones (see also Ben-Shakhar & Elaad, 2002, Honts, 2004, National Research Council, 2003, and Raskin, 1988).

Lack of Applicability

Perhaps the main criticism is that the GKT often cannot be used. For example, the test cannot be carried out if the suspect is not claiming lack of knowledge. The most common example here is an alleged sexual assault in which the witness claims that force was used and the suspect admits the sexual acts but claims that it was consensual. Similar problems arise in cases where several suspects admit to having been involved in the crime, but all of them deny having been the principal player (Raskin, 1988); or when someone claims to have been a witness to the crime rather than having been involved in commissioning it (Honts, 2004).

The application opportunities of the GKT are not just restricted because a suspect is not claiming lack of knowledge. Another restriction is that a GKT can only be carried out when the examiner knows some crime facts about which he or she can formulate test questions (but see Box 12.1). In several criminal investigations such crime facts are unknown. Podlesny (1993) reviewed criminal case files of the American Federal Bureau of Investigation (FBI) and found that in only 13.1% of the cases where polygraph examinations were used, could a GKT have been carried out.

Lykken (1998), however, claims that the GKT can be used in more cases than the FBI figures suggest. He pointed out that at present FBI investigators are not trained to search fresh crime scenes for usable GKT items. If they were so trained, the test could be used more often. He made a comparison with the search for fingerprints and stated that "Had Podlesny (*who analysed the FBI case files, just mentioned*) been working at Scotland Yard in 1900 at the time of the introduction of the Galton–Henry system of fingerprint identification, it is likely that he would also have found very few cases in the records of the Yard that included fingerprints of suspects." (Lykken, 1998, p. 305) (the information in italics has been added by me).

Moreover, I discussed in Chapter 11 that concern-based polygraph examinations are sometimes used for screening applicants or employees, or for checking on sex offenders. I also explained that in such situations the questions asked during the polygraph examination are not about specific events (e.g., "Did you tamper with the computer?") because there is no specific event that the examiner can refer to. Instead, non-specific questions are asked such as "In the last five months, have you ever provided classified information to an unauthorised person?" GKT examinations cannot be conducted in non-specific events cases.

Box 12.1 SPOT the unknown evidence

When critical details that can be used in a GKT are not available, a 'Searching Peak Of Tension' (SPOT) test is often carried out in Japan (Nakayama, 2002). The test looks identical to a GKT. Thus, in a murder case an examinee could be asked questions such as "When was he killed?", "Where was he killed?", "How was he killed?", and "Where is the corpse?" with each question followed by five answer alternatives. The difference with a GKT is that the correct answers to the questions are unknown and that it thus could be that a correct answer is not in the list of alternatives. If an examinee shows a distinctive physiological response to a certain item, the police could use this information to search for further evidence. For example, if the examinee shows a distinctive response to the item "lake" when asked where the corpse can be found, the police can elect to drag the lake.

Theoretical Concerns

The theoretical foundation of the GKT receives more support amongst scholars than the theoretical foundation of the concern-based polygraph tests discussed in Chapter 11 (Fiedler, Schmid, & Stahl, 2002; Iacono & Lykken, 1997; National Research Council, 2003). Yet, there are some concerns. The National Research Council's main concern is that reactions to familiar, personally significant stimuli and unfamiliar stimuli should be thought of as a continuum rather than a dichotomy (as suggested by GKT polygraphers). That is, suppose that the murderer used a revolver and suppose that the innocent examinee owns an unregistered pistol. That examinee might show responses to questions that mention handguns among the alternatives, even when he has no concealed knowledge about the murder weapon.

Honts' (2004) main concern is related to the memory of guilty suspects. He argued that there is no scientific approach to predict what elements of a crime scene culprits are likely to remember and thus recognise during the test. For orienting reflexes to occur it is essential that culprits recognise the critical details during the test. Honts refers to eyewitness testimony research that shows that people generally have problems with accurately remembering details of the scene of crime, and that people's memory can be distorted by the remarks and comments of other people.[3]

Difficulty in Formulating Proper GKT Questions

Selecting appropriate questions for a GKT is not an easy task. For a GKT to work, guilty examinees (i.e., culprits) should know the correct answers to the questions, otherwise they have no concealed knowledge, and innocent examinees should not be able to figure out what the correct answers are, otherwise they will be mistakenly accused of having

[3] The bulk of eyewitness testimony research, however, relates to how *witnesses* of crime perceive and remember events, and this may differ from how *culprits* perceive and remember events. See Christianson (2007) for an edited volume about the memory of offenders.

concealed knowledge. Culprits do not always know the correct answers. They may not have perceived the details the examiner is asking about, or may have forgotten them by the time the test takes place. For example, when the examiner asks a question about the colour of the shirt the man who was shot was wearing, the culprit may simply have forgotten the colour of that shirt or may never have noticed the colour.

In addition, orienting reflexes become stronger if the significance level of critical items increases. Thus, critical details about personally more significant information (i.e., examinee's own date of birth) result in stronger orienting reflexes than information about personally less significant information (i.e., a friend's date of birth), and orienting reflexes become stronger if the stakes are getting higher. Examiners thus should select critical items that they believe are likely to be remembered by and are significant to the culprit. Japanese examiners are trained in how to do this (see Nakayama, 2002, for some details). Selecting the appropriate critical items often involves visiting the crime scene, reading the case files, and, if possible, talking to victims or witnesses (Nakayama, 2002). However, as pointed out in the "theoretical concerns" section, the examiner cannot know what the culprit will remember. Indeed, although in Japan typically seven questions are asked in a GKT, it rarely happens that the suspect responds positively to all seven critical items (Nakayama, 2002).

One aspect that influences the culprit's memory is the time lapse between committing the crime and conducting the GKT. The longer the period between the crime and the GKT, the more likely it is that the culprit has forgotten certain details. GKT examinations should therefore be carried out as quickly as possible. Also, examiners may face particular difficulty in formulating appropriate questions for certain offenders. For example, serial offenders may have difficulty in recognising details related to the specific incident under investigation (Nakayama, 2002). They may have committed so many crimes that they forgot exactly what happened during one particular crime they are asked about. With serial offenders, questions should be asked about items that are striking and uniquely associated with the crime under investigation (Nakayama, 2002), but this may be difficult to achieve.

Once questions have been formulated, the task of selecting appropriate filler alternatives raises further difficulties. For example, if the content of filler alternatives is highly similar to the correct alternative, orienting responses will not discriminate between them (Ben-Shakhar, Gati, & Salamon, 1995; Nakayama, 2002). This would be harmful in demonstrating concealed information in culprits. Another possibility is that different alternatives elicit different responses in examinees because of their nature (e.g., intrapersonal differences). For example, it could be that the alternative "black bra" is more arousing, and thus leads to more distinctive physiological responses, than the alternatives "white bra", "grey bra" or "beige bra" (in a case where the examinee is a suspect in a rape case). It is also possible that some alternatives will evoke more distinctive responses because they sound more plausible. For example, it may sound more plausible that a burglar went into a house via an open window than by smashing the front window or forcing the front door. The alternative "open window" may therefore evoke in an innocent person a more distinctive physiological response than the alternatives about smashing a window or forcing the front door. If in these examples the alternatives "black bra" and "open window" were the correct alternatives, the physiological responses could be misinterpreted as orienting reflexes.

Many of these problems can be minimised by pre-testing the set of alternatives on known innocent people. They should give similar responses to all alternatives. Another set of filler alternatives needs to be chosen if this is not the case (Lykken, 1998). However, Honts (2004) pointed out that testing for a bias may be a daunting task in a field case because a substantial number of mock suspects need to be recruited and tested. Even without testing for a bias, setting up a GKT is time consuming as designing a test typically takes at least two or three days (Nakayama, 2002).

Leakage of GKT Items

Once the critical items about which the GKT questions will be formulated have been identified, it is important not to leak them to examinees. If culprits receive information about the critical items prior to a GKT, they can claim that this information rather than their presence at the scene of crime is the reason for their orienting responses during the GKT. If innocent suspects receive information about the critical items prior to the GKT, they too can produce orienting reflexes during the test. In other words, a GKT cannot be carried out when leakage of the GKT items has occurred.

Information about the critical items could have been made available through the media, attorneys, or investigators. Regarding the latter, in many countries the police must inform suspects of the crime facts directly after arrest, and this could include information that makes suspects aware of critical GKT items. For example, Japan is such a country where the police must inform the suspects about crime facts after their arrest, and therefore GKT examinations take place prior to the arrest of the suspect (Nakayama, 2002). Even in countries where it is not necessary to inform suspects of the crime facts, investigators often provide such information during the interview because it is seen as a potential method of extracting a confession (Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Leo, 1996a). Many suspects therefore could be aware of the critical items after being interviewed (Ben-Shakhar *et al.*, 2002).

To address these issues, examiners in Japan discuss with the examinee during the pre-interview phase of a GKT examination what knowledge he or she has about the crime under investigation. Questions about which an examinee indicates knowledge are replaced. However, innocent suspects could have acquired knowledge via leakage without being aware of it, and may therefore incriminate themselves during the test (Ben-Shakhar *et al.*, 2002).

Countermeasures

As I discussed earlier in this book, countermeasures are deliberate attempts by examinees to influence the verdicts of lie detection experts. For example, in GKT examinations, examinees could bite on their tongue, press their toes, or think about a frightening event each time a filler alternative is mentioned. The result is that the correct alternatives are less likely to stand out, and so guilty examinees could pass the test. Research has shown that countermeasures could be successfully used in GKT examinations (see Ben-Shakhar & Elaad, 2002, and Honts & Amato, 2002, for reviews). For example, Ben-Shakhar and Dolev (1996) instructed their Israeli participants to think about an emotional experience in the past when confronted with filler alternatives during the GKT. Some scenarios were suggested such as receiving an alarm about an approaching Scud missile attack during the First Gulf War. Those instructions reduced the accuracy of the GKT.

ACCURACY OF GKT POLYGRAPH EXAMINATIONS

To test the accuracy of GKT polygraph examinations both laboratory and field studies have been carried out. I already explained in previous chapters what the differences are between laboratory and field studies and which advantages and disadvantages are associated with each type of test. I will discuss the accuracy of GKT examinations in laboratory studies first, followed by the accuracy rates found in field studies.

In GKT examinations four outcomes are possible.[4] The *guilty suspect* (i.e., culprit) may show orienting responses to the critical alternatives

[4] In contrast to the CQT (Chapter 11), the GKT typically does not have an inconclusive category.

but not to the filler alternatives. In this case, the suspect fails the test, which is the correct decision. It is also possible that a guilty suspect shows similar responses to all alternatives, or shows a specific response to a filler alternative. In this case, the suspect will pass the test, which is an incorrect decision.

An *innocent suspect* may show similar responses to all alternatives or a specific response to a filler alternative. In this case, the suspect will pass the test, which is the correct decision. However, an innocent suspect may show a specific response to the critical item. This time the suspect fails the test, which is an incorrect decision.

GKT Accuracy Rates: Laboratory Studies

Table 12.2 shows the six reviews of laboratory studies that I am aware of that have been published in English. In GKT laboratory studies different paradigms were used, including the mock crime paradigm. In such studies, liars committed a mock crime and were instructed to conceal their knowledge about the crime in the subsequent GKT. Truth tellers did not commit the mock crime and therefore did not possess knowledge about the crime when asked about it during the GKT. Results of GKT laboratory studies show that the vast majority of guilty suspects were correctly classified (76% to 88%) but also that substantial numbers of guilty examinees (12% to 24%) passed the test. Most reviews show that GKT examinations are very accurate at pinpointing innocent examinees (94% to 99% accuracy rates), with hardly any innocent examinees (between 1% and 6%) falsely accused. However, MacLaren's (2001) review shows a different picture with fewer innocent examinees correctly (83%) and more of them incorrectly (17%) classified.

Table 12.2 Results of laboratory study reviews examining the accuracy of GKT polygraph testing

	Guilty condition		Innocent condition	
	Guilty	Innocent	Guilty	Innocent
Ben-Shakhar & Furedy (1990, $n = 10$)	84%	16%	6%	94%
Ekman (1985/2001, $n = 6$)	78%	22%	5%	95%
Elaad (1998, $n = 5$)	81%	19%	4%	96%
Honts (1995, $n = 5$)	86%	14%	1%	99%
Lykken (1998, $n = 8$)	88%	12%	3%	97%
MacLaren (2001, $n = 22$)	76%	24%	17%	83%

Note: n = number of studies reviewed

The difference between MacLaren's review and the other reviews is that MacLaren included many more studies. However, even MacLaren used selection criteria. He started with 78 studies but reported the results of only 22 studies. Ben-Shakhar and Elaad (2003) also carried out a review of GKT laboratory studies and did not leave out any of the studies they had collected. They reviewed a total of 80 studies. They did not report accuracy rates (hence, the review is not reported in Table 12.2) but calculated effect sizes, d. The higher the effect size, the better the GKT discriminated between truth tellers and liars. As I explained earlier in this book, effect sizes of $d > .8$ are considered to be large. One benefit of using effect sizes is that differences between groups of studies can be calculated, which is what Ben-Shakhar and Elaad have done. The average effect size across all 80 studies was $d = 1.54$. When they examined only the mock crime studies, an effect size of $d = 2.09$ was obtained. These are large effects. For example, effect sizes for nonverbal cues to deceit, reported in Chapter 3, are typically around $d = .20$. In fact, effect sizes of $d > .70$ are not very often reported in behavioural sciences (Ben-Shakhar & Elaad, 2003; Bond & DePaulo, 2006). Ben-Shakhar and Elaad (2003) further found that the effect was stronger in studies where participants were motivated to perform well ($d = 1.84$) than in studies where no motivation was given ($d = 1.36$), which they claim to be consistent with the theoretical notion that if the level of significance of critical items increases, the orienting reflex increases. Finally, GKT examinations based on more than five questions produced a larger effect size ($d = 2.35$) than GKT examinations based on five questions or less ($d = 1.29$).

GKT Accuracy Rates: Field Studies

Only two GKT field studies have been published to date, both of which were conducted in Israel (see Table 12.3). Results were extraordinary regarding innocent examinees, with 98% and 94% of them being correctly classified.[5] The results for guilty examinees were less impressive, and also differed between the two studies. In one study a reasonable percentage (76%) of guilty examinees were correctly classified, but in the other study the accuracy rate (42%) of guilty examinees was very low. However, both field studies had limitations. For example, the number of questions asked was low (around two questions per GKT on average), and, as Ben-Skakhar and Elaad's (2003) review revealed, GKT

[5] Those accuracy rates suggest that the accuracy rates in GKT laboratory studies regarding innocent examinees reported by MacLaren (2001) were exceptionally low rather than that the accuracy rates reported by the others being exceptionally high (Table 12.2).

Table 12.3 Results of field studies examining the accuracy of GKT polygraph testing

	Guilty condition		Innocent condition	
	Guilty	Innocent	Guilty	Innocent
Elaad (1990)	42%	58%	2%	98%
Elaad *et al.* (1992)	76%	24%	6%	94%

examinations do increase in accuracy in correlation with the number of questions that are asked.

The laboratory and field studies both show that the GKT is vulnerable to false-negative errors, which is the tendency to classify guilty examinees as innocent. Although such errors are far from desirable, they are in criminal investigations generally perceived as less serious than false-positive errors (Chapter 11), which is the tendency to classify innocent examinees as guilty. It is not surprising that the GKT is vulnerable to false-negative errors. These occur when the suspect has forgotten or never knew the details about which the examiner asks. By the same token it is unsurprising that field studies are more vulnerable to false-negative errors than laboratory studies. In laboratory settings researchers can increase the likelihood that guilty examinees remember the critical details, for example by carrying out the test directly after the guilty examinee has been given the critical information. In real life it may be considerably more difficult to select items that culprits are likely to remember. In that respect, laboratory research masks this problem associated with the GKT.

Box 12.2 Reaction times in GKT examinations

Reaction times in responding to critical items and filler alternatives are increasingly used in deception studies that are based on the GKT paradigm. In a popular research setting participants read critical items and filler alternatives that appear on computer screens. They are requested to deny knowledge of the items as quickly as they can by pressing a "No" button on their keyboards. They are also shown a third set of items to which they have to respond "Yes" (otherwise they could simply press "No" all the time). Researchers compare the reaction times in answering "No" to the critical items and filler alternatives. Those studies, together with the studies that measure reaction times in different, non-GKT, research settings, are included in Chapter 3 (Appendix 3.1 under the heading *latency period*).

I have doubts about using reaction times as a measurement of deceit in real-life settings. First, relying on a single measure (reaction time) is relying on Pinocchio's growing nose which will not work as I explained in previous chapters. Second, I expect this single measurement test to be highly sensitive to countermeasures. Third, as I already mentioned, selecting critical and filler alternatives is a difficult task for GKT examiners. If they are going to measure reaction times they will have an additional problem to address. Not all critical items and filler alternatives will be understood equally quickly. The longer it takes to fully grasp the question, the longer the reaction time will be. This will affect the test scores, particularly because reaction times are typically measured in 1/100s of seconds. Thus, in a reaction time test, an examinee may be accused of lying simply because he or she did not understand the critical item as quickly as the filler alternatives.

Moreover, I can envisage that university students in laboratory settings are keen to carry out reaction time tasks but I am not convinced that other people, for example suspects in police interviews, will be equally dedicated to doing so. Neither do I think that such tasks will suit all people equally well. Having experience in using computers probably affects people's reaction times. In addition, reaction times cannot be used with people who are illiterate, have difficulty with reading, are dyslexic, have poor eyesight, and so on. Finally, suppose that someone fails the reaction time test. How likely is it that the he or she will accept the test outcome? It will not be difficult for them to give various harmless reasons that can explain why they failed the test.

LEGAL IMPLICATIONS

In some previous chapters I have discussed the implications for a particular lie detection tool in light of the findings that I have highlighted, and so I will now discuss what I think the implications are of this chapter's findings for the use of GKT polygraph outcomes as scientific evidence in criminal courts. Again, I will use the set of guidelines provided by the United States Supreme Court for admitting expert scientific evidence in the federal (American) courts (*Daubert v. Merrel Dow Pharmaceuticals, Inc.*, 1993). Table 12.4 summarises my verdict.

Table 12.4 Answers to the five *Daubert* questions for GKT assessments

	GKT laboratory	GKT field
(1) Is the scientific hypothesis testable?	Yes	Problematic, but is it necessary?
(2) Has the proposition been tested?	Yes	Rarely
(3) Is there a known error rate?	Yes	Yes, too high for guilty examinees
(4) Has the hypothesis and/or technique been subjected to peer review and publication?	Yes	Yes, but only two studies published
(5) Is the theory upon which the hypothesis and/or technique is based generally accepted in the appropriate scientific community?	Yes	Yes

Question 1: Is the Scientific Hypothesis Testable?

The underlying rationale of the GKT, the orienting reflex, is easily testable in laboratory studies but more problematic in field studies because it is difficult to establish the ground truth in such cases. In defence of the GKT, it could be argued that testing the orienting response in a field setting is not necessary. It would be necessary if there are theoretical reasons why orienting responses to critical items in laboratory situations would differ from those in real life, but there are no such reasons, because orienting responses do not depend on those factors in which laboratory studies and field studies differ, such as stress or anxiety (Kugelmass & Lieblich, 1966, but see also Carmel, Dayan, Naveh, Raveh, & Ben-Shakhar, 2003).

Question 2: Has the Proposition been Tested?

The proposition has been tested in laboratory research but only twice in field research.

Question 3: Is there a Known Error Rate?

The error rate is known in laboratory research, and is acceptable for innocent examinees (in most literature reviews less than 6%) but is somewhat high for guilty examinees (between 12% and 24% according to the literature reviews). The known error rate in field studies is acceptable for innocent examinees (2% to 6%), but high for guilty examinees (ranging from 24% to 58%). The latter accuracy rates do not pass the "beyond reasonable doubt" criterion, often used as a standard in criminal courts.

Question 4: Has the Hypothesis and/or Technique been Subjected to Peer Review and Publication?

Yes, both laboratory and field research GKT research has been subjected to peer review, although only two field studies have been published.

Question 5: Is the Theory upon which the Hypothesis and/or Technique is Based Generally Accepted in the Appropriate Scientific Community?

In their survey of scientific opinion about the GKT, Iacono and Lykken (1997) consulted members of the American Society of Psychophysiological Research (who can be considered as experts) and fellows of the American Psychological Association (Division 1, General Psychology) about GKT polygraph testing. The experts and fellows were asked to consider whether the GKT is based on scientifically sound psychological principles or theory. Both experts and fellows expressed similar opinions, and most (around 75%) agreed that the GKT is based on a sound rationale. In other words, GKT polygraph examinations are generally accepted by the relevant scientific community.

Overall Verdict

In my view, the GKT does not meet the *Daubert* guidelines for admitting expert scientific evidence in criminal courts. The main problem is that the known error rate is too high, albeit only for guilty suspects. One could argue that the two field studies were limited. On average only two questions were asked in each GKT examination, and GKT examinations become less accurate when only a few questions are asked. In that respect a future field study in which more questions are asked may yield more positive results for GKT testing, making it more likely that GKT examinations will past the *Daubert* test.

EVENT-RELATED POTENTIALS: THE P300 BRAIN WAVE

Event-related potentials are brain waves that are recorded via electroencephalograms (EEGs) (Rosenfeld, 2002). One of these waves, the P300, is of interest to deception researchers because it occurs in response to personally significant stimuli, and can thus be seen as an orienting reflex. The brain waves are called P300s because they have their peak typically after 300 to 1000 milliseconds from the stimulus onset. EEG-P300 deception research was initiated in Japan, and the first article was published (in Japanese) in 1986 (Nakayama, 2002). This type of deception research continues to flourish in Japan but most of these studies are published in Japanese. The first EEG-P300 deception study in English appeared in 1988 (Rosenfeld, 2002) and the number of publications is ever increasing.[6]

Several of those studies in English (all laboratory studies) reported accuracy rates and those are presented in Table 12.5. The findings demonstrate that P300 brain waves can be used to detect deception. For guilty participants, 51% to 100% were correctly classified (82.29% accuracy rate) and between 0% and 49% (16.21% accuracy rate) were incorrectly considered innocent. For innocent participants, 72% to 100% were correctly classified (87.50% accuracy rate) and between 0% and 24% (8.75% accuracy rate) were incorrectly considered guilty. These accuracy rates are similar to those found with traditional polygraph measures (see Table 12.2 and also National Research Council, 2003), but Japanese researchers believe that P300 measures may classify truth tellers and liars somewhat more accurately than the traditional polygraph measures (Hira, Furumitsu, & Nakayama, 2006; Nakayama, 2002).

One comment is relevant here. The only difference between traditional GKT polygraph examinations and EEG-P300 examinations utilising a GKT paradigm is how the orienting response is measured (with EDA, blood pressure and respiration or with P300 waves). The problems with GKT examinations, however, are not related to measuring the orienting response. The problems are associated, for example, with the limited applicability of the test; selecting critical items that culprits are likely to remember and innocent suspects cannot easily figure out; and the leakage of critical information. Using P300 measurements rather than the traditional polygraph measurements does not solve any of these problems.

[6] Abootalebi, Moradi, & Khalilzadeh, 2006; Allan & Iacono, 1997; Allen, Iacono, & Danielson, 1992; Boaz, Perry, Raney, Fischler, & Shuman, 1991; Ellwanger, Rosenfeld, Hannkin, & Sweet, 1999; Ellwanger, Rosenfeld, Sweet, & Bhatt, 1996; Farwell & Donchin, 1991; Farwell & Smith, 2001; Johnson, Barnhardt, & Zhu, 2003, 2004, 2005; Johnson & Rosenfeld, 1992; Rosenfeld, Angell, Johnson, & Qian, 1991; Rosenfeld, Biroschak, & Furedy, 2006; Rosenfeld, Cantwell, Nasman, Wojdac, Ivanov, & Mazzeri, 1988; Rosenfeld, Ellwanger, Nolan, Wu, Bermann, & Sweet, 1999; Rosenfeld, Ellwanger, & Sweet, 1995; Rosenfeld, Rao, Soskins, & Miller, 2003; Rosenfeld, Reinhart, Bhatt, Ellwanger, Gora, Sekera, & Sweet, 1998; Rosenfeld, Shue, & Singer, 2007; Rosenfeld, Soskins, Bosh, & Ryan, 2004). See Bashore and Rapp (1993) and Rosenfeld (1995, 2002) for reviews.

Table 12.5 Results of laboratory studies examining the accuracy of the GKT using EEGs measuring P300 brain waves

	Guilty condition (%)			Innocent condition (%)		
Authors	Guilty	Innocent	Inconclusive	Guilty	Innocent	Inconclusive
Abootalebi *et al.* (2006)[1]	68	32		16	84	
Allen & Iacono (1997)[2]	87	0	13	0	72	28
Allen *et al.* (1992)	90	10		4	96	
Boaz *et al.* (1991)[3]	76	24		24	76	
Ellwanger *et al.* (1996)[4]	73	27		0	100	
Farwell & Donchin (1991)[5]	92	0	8	0	83	17
Johnson & Rosenfeld (1992)	76	24		21	79	
Rosenfeld, Angell *et al.* (1991)[6]	92	8		13	87	
Rosenfeld, Biroschak *et al.* (2006)[7]	90	10		18	82	
Rosenfeld, Cantwell *et al.* (1988)	100	0		0	100	
Rosenfeld, Ellwanger *et al.* (1999)	87	13		0	100	
Rosenfeld, Ellwanger *et al.* (1995)[8]	88	12				
Rosenfeld, Shue *et al.* (2007)[9]	51	49				
Rosenfeld, Soskins *et al.* (2004)[10]	82	18		9%	91	
Total	**82.29**	**16.21**		**8.75**	**87.50**	

Notes: [1]Average across three analyses; [2]Average across all three conditions; [3]Average across all three groups of participants; [4]Table 5, t-test results; [5]I combined Experiments 1 and 2 because only four participants took part in Experiment 2; [6]Without delay condition; [7]Meaningful stimulus condition; [8]Average across all target words; [9]All conditions combined; [10]Control condition only.

Two studies clearly demonstrated the vulnerability of EEG-P300 lie detection. In all EEG-P300 deception studies presented in Table 12.5 guilty examinees were tested immediately after they had been told which information to conceal. These are optimum testing conditions because it makes it likely that the guilty examinees will remember the information. In one experiment, guilty examinees were either tested under such optimum conditions or with a 7–14 day delay. Although virtually all guilty examinees (12 out of 13) were detected in the optimum condition (Table 12.5), only three out of eight guilty examinees were correctly diagnosed in the delay condition (Rosenfeld, Angell, Johnson, & Qian, 1991). In other words, EEG-P300 examinations utilising a GKT paradigm are, like GKT polygraph examinations, vulnerable to guilty suspects forgetting the information they are asked about.

Moreover, in many of these laboratory studies the information that participants were asked to conceal was personally meaningful (e.g., participants' names). As mentioned above, personally meaningful information, because of its psychological significance, will evoke larger P300 brain waves than less meaningful information, and this makes lie detection via measuring P300 brainwaves easier when the guilty knowledge consists of highly meaningful rather than less meaningful information. Indeed, in one study, concealment of highly meaningful information (the participant's own name) could be detected with 90% accuracy (Table 12.5), but concealment of less meaningful information (the experimenter's name) could only be detected with 40% accuracy (Rosenfeld, Biroschak, & Furedy, 2006). All participants remembered the name of the experimenter, thus the low accuracy rate did not occur because they had forgotten the information; rather the experimenter's name was not personally significant enough to the participants to evoke clear distinguishable P300 brain waves. EEG-P300 lie detection tests can thus only be used with critical details that are highly meaningful to the participants.

One could further argue that EEG-P300 examinations using a GKT paradigm have additional disadvantages compared to GKT polygraph examinations. EEG examinations are highly intrusive, as examinees have to wear an uncomfortable cap containing as many as 128 electrodes. Such caps are not really suitable yet for using outside the laboratory, which is one reason why practitioners do not measure P300 brain waves to date (Nakayama, 2002). In addition, the P300 brain wave is typically a very small signal, and so to measure the P300 brain waves reliably, many trials are needed. For example, Rosenfeld, Shue, and Singer (2007) showed their participants the critical items and filler alternatives 30 times. In other words, EEG-P300 examinations are considerably more difficult to conduct than polygraph examinations.

Initially it was suggested that P300 waves are more resistant to countermeasures than the traditional polygraph measures (Ben-Shakhar & Elaad, 2002; Iacono, 2001; MacLaren, 2001). Using indices of orienting responses that are resistant to countermeasures would substantially improve the value of the GKT in real life. However, unfortunately, P300 waves are vulnerable to countermeasures. In a GKT experiment, one group of guilty participants were asked to employ countermeasures, such as pressing a finger or wiggling a toe each time they saw an alternative other than the critical item, whereas no instructions were given to another group of guilty participants. Meanwhile the participants' P300 waves were measured. The untrained guilty participants were detected with 82% accuracy (Table 12.5), but the participants trained in countermeasures were detected with only 18% accuracy (Rosenfeld, Soskins, Bosh, & Ryan, 2004).

Box 12.3 MERMER: the P300+ approach

Lawrence Farwell has examined whether a compilation of brain waves could accurately distinguish truth tellers from liars. This brain wave pattern, with the abbreviation MERMER, consists of P300 waves and other brain waves. In their experiment, Farwell and Smith (2001) carried out a GKT with six participants, three with guilty knowledge and three without guilty knowledge. They could correctly classify all six participants. Farwell describes his MERMER technique, which is patented (Ford, 2006), as *Brain Fingerprinting*, and promotes it on his website, www.brainwavescience.com (see also Chapter 1). Remarkably, in their article, the only published article with the MERMER technique that I am aware of, they refer to polygraph testing and claim that "Brain MERMER testing . . .. has almost nothing in common with lie detection or polygraphy" (Farwell & Smith, 2001, p. 142). This phrase is misleading. Although it has nothing to do with concerned-based polygraph tests, which are discussed in Chapter 11, it has much in common with GKT polygraphy. The only difference between the MERMER technique and traditional GKT polygraph testing is how the orienting response is measured. As I just explained in the main text, measuring the orienting response is not the problem in GKT examinations. Rather, the limited applicability of the test, selecting critical items, and the leaking of GKT items are the concerns. MERMER is equally vulnerable to these concerns as traditional GKT examinations. Another problem with the GKT is that examinees could use countermeasures to influence the test outcomes. Since the MERMER brain wave pattern includes P300 waves, it may also be vulnerable to countermeasures. Intriguingly, Farwell and Smith (2001) discuss the limited applicability of their technique as the only limitation and do not discuss or refer to the relevant GKT polygraph literature.

EVALUATION OF THE GKT

The rationale of the GKT is that guilty examinees show enhanced orienting reflexes when recognising crucial details of the crime. Problems associated with the GKT are the limited applicability of the test; difficulties in formulating questions to which the guilty examinee knows the correct answer and to which the innocent examinee cannot figure out the correct answer; the leakage of critical information to examinees; and its vulnerability to countermeasures.

This chapter demonstrated that the GKT has a sound rationale and that it can classify liars, but in particular truth tellers, well above the level of chance. However, I do not think that the test outcomes should be used as evidence in criminal courts. For this purpose the error rate in pinpointing liars (i.e., false negative errors) is too high. False negative errors result in criminals going unconvicted where criminal courts rely heavily on the GKT outcome. Moreover, if GKT outcomes are used as evidence in criminal courts, investigators may attempt to solve cases via conducting a GKT rather than via gathering evidence that conclusively links the suspect to the crime. I think this risk is considerable because conducting a GKT is probably less time consuming than the often painstaking practice of gathering evidence. Also, if GKT outcomes were accepted as evidence in criminal courts, guilty suspects too will recognise the importance of a GKT outcome, which may result in them attempting to beat the test. People are capable of doing this when applying certain countermeasures.

Given the GKT's strong theoretical rationale, its ability to detect truths and lies above the level of chance, and the safeguard it provides for innocent suspects, I encourage practitioners to use the GKT. However, to introduce a GKT involves more than just buying a polygraph, and training people how to use the equipment. Setting up a good GKT involves selecting appropriate critical items and filler alternatives. This is not a straightforward task and people need to be properly trained in how to do this. Also the issue of leaking the correct GKT alternatives to examinees prior to a GKT examination requires careful attention. In this respect, much can be learned from the Japanese who have built up considerable experience in conducting GKT examinations.

Much of the current GKT research is related to measuring the orienting response in new ways, for example, via P300 brain waves. Whether this will improve the accuracy of the GKT remains to be seen, because adequately measuring the orienting response is not the problem in GKT examinations. One aspect of the GKT that needs improvement is that liars sometimes remain undetected. The reason for this is that they do not recognise the critical items on the test, and consequently, the orienting response does not occur. In other words, the lack of presence of an orienting response is the problem, not the way an orienting response is measured once it occurs. Therefore, to improve the accuracy of the GKT, rather than measuring the orienting response in a different way, more will be gained from increasing our understanding of (i) which type of critical items guilty suspects are likely to remember, and (ii) which factors, both whilst committing the crime and after the crime, affect the recall of critical items during the test. Those aspects are addressed in eyewitness research and perhaps the extensive eyewitness testimony literature can give examiners useful ideas about this. See Christianson (2007) and Toglia, Read, Ross, & Lindsay (2006) for an overview of this literature.

Apart from the difficulty in selecting critical items, another problem of the GKT is that it cannot be used in all situations. For example, to select critical items for the test, the examiner needs information about what happened during the event under investigation, and suspects should deny having this knowledge. Situations where no information is available, or where the examinee claims other reasons for having crime-related knowledge ("I witnessed the crime, but I did not commit the crime") are thus inappropriate. Similarly a GKT cannot be used when the critical information is leaked to the examinee via the media, attorneys, criminal investigators, and so on, because this leakage may result in orienting responses. To measure the orienting reflex in alternative ways will not solve these issues and therefore will not widen the GKT's applicability.

I do not wish to discourage researchers from examining alternative ways to measure the orienting response. I find this useful for at least two reasons. First, yet another weakness of the GKT is that it is vulnerable to countermeasures. Therefore, an important focus of such research should be to examine whether the orienting response can be measured in ways that are beyond the control of the examinees and therefore not sensitive to countermeasures. Second, the equipment used in brain waves research looks impressive, as do the brain wave charts that the equipment produces. It all probably looks more impressive than the traditional polygraph. I have the impression that many government officials are inclined to support lie detection tools that use modern, sophisticated looking equipment and, therefore, I hope that measuring the orienting reflex with such modern equipment will enhance the appeal of the GKT amongst government officials.

CHAPTER 13

Physiological Lie Detection: functional Magnetic Resonance Imaging (fMRI)

This short chapter discusses research related to lie detection via measuring activity in brain structures and areas, which is the most recent development in physiological lie detection. Measuring brain structure and area activity sounds intriguing. Some researchers describe it as examining "directly the organ that produces the lie" (Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003, p. 830) and others refer to it as gaining "direct access to the seat of a person's thoughts, feelings, intentions and knowledge" (Wolpe, Foster, & Langleben, 2005). In this respect, some have argued that researchers' ability to read a person's mind fuels ethical worries (Happel, 2005; Pearson, 2006; Wild, 2005; Wolpe *et al.*, 2005, see also *The Observer*, 20 March 2005). However, it is the potential of brain structure and area activity measurements to discriminate between truths and lies that I will focus on in this chapter.

Activity in brain structures and areas is associated with changes in blood flow and oxygen consumption that can be measured with a functional magnetic resonance imaging (fMRI) scanner. These scanners are primarily used in hospitals, for example, to detect brain tumours or injuries, but nowadays are also used by scientists in their attempts to detect lies. fMRI scanning has its disadvantages. The scanners are very expensive, as are the individual fMRI scans. Moreover, fMRI scanning is time consuming. A scan for lie detection purposes can last one hour and the complex data analyses can take several additional hours. In addition, being a participant in an fMRI lie detection test is an uncomfortable experience. The inside of an fMRI scanner is narrow and dark, with only a sliver of the world visible in a tilted mirror above the eyes. Participants have to lay flat on their back, are not allowed to move, and their heads are immobilised by pillows and straps. The machine is also very noisy, even when the participant wears earplugs (Silberman, 2006). Also, the machine uses a very strong magnet and so any kind of metallic objects, such as piercings, need to be removed prior to entering the scanner. Any ferrous objects within a person's body, for example fractured bones that have been repaired with screws, wires or metal plates, or a bullet or shrapnel that has lodged in the body, could result in fatality, so people with such ferrous objects within their body should not participate; pregnant women and people with claustrophobia are advised not to participate either. Finally, verbal communication with the outside world is not really possible during the examination. Participants tell the truth or lie by pressing buttons and can use a panic button if they need to. This all makes the fMRI lie detection tool rather difficult to apply in real life, and, in my view, only worthwhile considering if it really works, that is, if fMRI scans are able to discriminate between truths and lies at accuracy levels that exceed levels that can be achieved with traditional methods that are easier to apply.

This chapter discusses the fMRI lie detection studies that have been published to date. This research is still in its initial stage and many aspects relevant to lie detection have yet not been addressed. However, these initial studies have already given valuable insights. First, they revealed that fMRI analyses can successfully discriminate between truths and lies but not at a level that exceeds some methods discussed in previous chapters. Second, some of the problems that lie detectors face when using other methods, such as the absence of a cue uniquely related to deception (e.g., the absence of a cue akin Pinocchio's growing nose) and difficulty in controlling for intrapersonal differences, also apply to lie detection via fMRI scanning. fMRI scans, however, can give us valuable insight into the theoretical understanding of the mental processes that are associated with lying.

fMRI LIE DETECTION RESEARCH

The first fMRI lie detection study was published in 2001 by Sean Spence and his collaborators, and to my knowledge 11 more studies have been published since then. All of these studies are summarised in Table 13.1.

Table 13.1 An overview of fMRI lie detection studies

Authors	Task. Lie or tell the truth about:	Results. Brain activities are associated with:
Davatzikos *et al.* (2005)	Having seen a play card	Inhibitory function, error monitoring
Ganis *et al.* (2003)	Autobiographical information	Inhibitory function, conflict response
Kozel *et al.* (2004a)	Under which object money was hidden	Inhibitory function, emotion regulation, multi tasking
Kozel *et al.* (2004b)	Under which object money was hidden	Inhibitory function, divided attention
Kozel *et al.* (2005)	Mock crime	Inhibitory function, high-order decision
Langleben *et al.* (2002)	Having seen a play card	Inhibitory function, conflict response
Langleben *et al.* (2005)	Having seen a play card	Inhibitory function, attentional and contextual working memory system
Lee *et al.* (2002)	Autobiographical information	Executive functions control
Mohamed *et al.* (2006)	Mock crime	Inhibitory function, anxiety, memory encoding and retrieval, linguistic processing
Nunez *et al.* (2005)	(Non)autobiographical information	Inhibitory function, conflict response, cognitive control
Phan *et al.* (2005)	Having seen a play card	Inhibitory function
Spence *et al.* (2001)	Autobiographical information	Inhibitory function

The publications provide detailed information about the structures and areas of the brain that were activated when lying, and about the psychological processes associated with these brain activities. I discuss only these psychological processes and have summarised them in Table 13.1.

All studies published so far are laboratory experiments where examinees told the truth and lied for the sake of the experiment. In some experiments, participants were promised money if they could lie successfully, in other studies no incentives were given. In all studies participants told the truth and lied by pressing different buttons or raising different fingers. I have summarised the experimental designs used in these studies in Table 13.1.

In the "play card paradigm" a GKT was employed (Chapter 12). Participants were given a play card prior to the fMRI test, and were shown this card amongst other play cards during the fMRI test. They were instructed to deny possession of each card they were shown, resulting in lying about the card they had in their possession and telling the truth about the other cards.

In the remaining three paradigms, questioning techniques that resemble the concern-based techniques were used (Chapter 11). In the "autobiographic paradigm" participants were asked questions about themselves ("Do you own a laptop?") or about their daily activities (e.g., "Did you make your bed today?"; "Have you taken a tablet today?") and were instructed to answer some questions truthfully and some questions deceptively. In the "hidden money paradigm" participants were shown under which two out of 10 objects US $50 bills were hidden. They were then shown each of the 10 objects and were asked each time whether the money was hidden under that particular object. They were instructed to tell the truth about the location of one US $50 bill but lie about the location of the other by suggesting it was hidden under a different object than it in fact was. In the "mock crime paradigm" participants took either a ring or a watch from a drawer and put it in a locker. During the test they were asked a series of questions about both the ring and watch (e.g., "Did you take the ring from the drawer?"; "Is the watch in your locker?") and were instructed to answer the questions about the ring and watch as if they had taken neither object.

In all those studies it was found that lying was associated with specific brain structure and area activity. In particular, areas associated with inhibition and conflict resolution were activated when lying, suggesting that it is the suppression of the truth (e.g., inhibition) and the conflict between knowing the truth and telling a lie that gave the lie away. In their review of the literature, Spence *et al.* (2004) reported that deception generally activates the higher centres of the brain which are typically associated with cognitive demand. In other words, at least in the studies conducted to date, participants found it mentally more difficult to lie than to tell the truth.

Despite some similarities in the findings between the different studies there were considerable differences. In each study a somewhat different brain activity pattern emerged as an indicator of deceit, which indicates that a pattern uniquely related to deception does not exist in the brain (Abe *et al.*, 2006). This is further emphasised by Ganis *et al.*'s (2003) study that revealed that telling spontaneous lies resulted in different brain structure and area activity than telling rehearsed lies. Moreover, issues such as whether someone feels strongly about the issue under investigation and the stakes of getting caught, influence brain structure and area activity (Ganis *et al.*, 2003). In other words, their study found that people's brain activity when lying depends on the situation (so-called intrapersonal differences, see previous chapters). Kozel and colleagues (2004a,b) further reported that in the same lie detection experiment different individuals showed somewhat different brain activity patterns (so-called interpersonal differences, see previous chapters).

In three studies it was reported the extent to which truths and lies could be correctly classified on the basis of brain structure and area activity (Davatzikos *et al.*, 2005; Kozel *et al.*, 2005; Langleben *et al.*, 2005). The accuracy rates ranged from 78% (Langleben *et al.*, 2005) to 93% (Kozel *et al.*, 2005). Both Langleben and Kozel have subsequently applied to patent their fMRI-based lie detection instrument and are seeking to commercialise it (Happel, 2005). Researchers from the Langleben group have also announced that "brain imaging is ready to detect terrorists" (Wild, 2005).

A CLOSE LOOK AT THE fMRI LIE DETECTION TOOL

Let's have a more detailed look at fMRI scanning and lie detection. First, the fMRI studies conducted to date were all laboratory studies. As I explained in Chapters 11 and 12, laboratory settings present optimum conditions for physiological lie detection because they mask some of the difficulties that lie detectors face in real life. A main difficulty in concern-based tests is formulating questions that are appropriate controls for the relevant questions. Thus, the issue is which control questions to ask in fMRI tests where the relevant questions are "Are you involved in terrorist activity?" or "On 26 January, did you shoot Julie Pearchap?" I explained in Chapter 11 that many scientists believe that appropriate control questions can never be asked in such situations because the stakes are almost inevitably higher when answering the relevant questions than when answering any control question.

Also a GKT paradigm is used in fMRI research but this paradigm has other limitations (Chapter 12). It can only be used in situations where some critical details about the examinee's alleged activities are known. In addition, a GKT can only work if guilty examinees actually recognise the critical information when it is presented to them during the examination, and when innocent examinees cannot figure out what the critical information is. Achieving this can be a daunting task. Finally, sometimes the critical information is leaked to examinees prior to the examination. If this happens, a GKT cannot be conducted. fMRI tests based on a GKT are equally vulnerable to these limitations as traditional GKT polygraph examinations. Second, the accuracy rates reported in laboratory polygraph examinations were also high (Chapters 11 and 12) and comparable with the accuracy rates of the fMRI studies. This raises doubts about the added value of fMRI examinations compared to traditional polygraph examinations. Third, if fMRI examinations were to be used in security threat investigations to detect potential terrorists or spies, accuracy levels would need to be very high, due to the low base rate of terrorists and spies in a given sample (Chapter 11). Suppose that a given sample of 1000 people contains 10 spies, and suppose that the fMRI test can discriminate between truths and lies with 90% accuracy. Such an fMRI test would detect most (nine) but not all spies. It would further incorrectly classify approximately 99 innocent people as spies (10% of 990 non-spies), and the fMRI test outcomes cannot further discriminate between the nine spies and the 99 innocent individuals. In other words, a lie detection tool with 90% accuracy is not accurate enough to detect a few terrorists and spies in a large sample.

Fourth, in the fMRI studies where accuracy rates were reported the researchers identified specific brain activity models that involved simultaneous activity of different areas of the brain, and could classify truths and lies with reasonable accuracy on the basis of these brain activity models. However, their findings demonstrate that their models work in their specific deception scenarios, but do not demonstrate that they will work in other deception scenarios. Their models may not work as accurately in other scenarios because different types of lies result in different brain activities (see above). Moreover, the participants in these studies were undergraduate students or "unmedicated adults" (Kozel *et al.*, 2005). There is no guarantee that their models will also work accurately with other types of individuals, such as psychopaths, pathological liars, people who use drugs, depressed individuals, and so on (see Yang *et al.*, 2005 for the brain structure and area activity of pathological liars). It could also be that jihadists who think that Americans and other Western people are infidels have a different frame of mind than the participants in the fMRI experiments carried out to date, making it unclear how their brain structure and area activity responds when they are lying (Bloche, quoted in Pearson, 2006). In other words, the research that convincingly demonstrates that fMRI scans can distinguish between truths and lies in different situations and with different types of individuals has not been published yet. Such evidence is required if one is to rely upon these tests to catch terrorists and other types of serious criminals.

Fifth, if fMRI lie detection tests were to be introduced in real-life settings, guilty examinees may try to beat these tests. For the fMRI test to be effective and thus safe to rely upon, it should be resistant to countermeasures. No fMRI countermeasures studies have been published to date, but preliminary data suggest that easy to learn countermeasures could reduce the accuracy of fMRI lie detection tests (Langleben, Dattilio, & Guthei, 2006; Wolpe *et al.*, 2005).[1]

EVALUATION OF THE fMRI LIE DETECTION TECHNIQUE

fMRI scanning techniques used in hospitals to detect brain tumours or injuries are now used by scientists to detect lies. Those scientists look at the brain structure and area activity during truth telling and lying. Some of these brain activities have been found to be associated with deception, particularly those related to inhibition and response conflict. In general terms, deception activates higher centres of the brain which suggests that, at least in the fMRI studies carried out to date, people find it somewhat more difficult to lie than to tell the truth. Although fMRI lie detection is a new area of research with only a few studies published to date, several limitations are already clear. For example, different people tested in the same situation revealed different patterns of brain structure and area activity when they lied (interpersonal differences) and the same person shows different patterns of brain structure and area activity when he or she lies in different situations (intrapersonal differences). This makes fMRI lie detection not fundamentally different from lie detection with a traditional polygraph and the limitations

[1] I doubt whether the scientists who asked for the patent are fully aware of the challenges that fMRI lie detection faces. For example, one member of the Langleben team was quoted as saying "A lie is always more complicated than the truth. . . You think a bit more and fMRI picks that up" (Wild, 2006). Lying, however, is not always more demanding than truth telling (Chapters 3 and 15).

associated with polygraph testing discussed in Chapters 11 and 12 also apply to fMRI lie detection.

fMRI tests are expensive, time consuming, and uncomfortable for examinees. I therefore argued in the beginning of this chapter that such tests are only worthy of introducing as a lie detection tool in real-life situations if they are more accurate than the alternative techniques available to date. So far, research has not yet shown that the fMRI technique does produce more accurate results than traditional polygraph testing, and I therefore do not recommend using such scans in real-life settings for lie detection purposes (see also Langleben, in press; Spence, in press).

I do encourage scientists to conduct fMRI lie detection research because I believe that valuable information can be learned from this technique. In particular, it could give us valuable insight into the theoretical understanding of the mental processes associated with deception. Such information could be important to develop new, more effective, lie detection techniques (National Research Council, 2003). In Chapter 15 I will elaborate on the sort of research activities I have in mind.

CHAPTER 14

Pitfalls: Why People Fail To Catch Liars

The final two chapters of this book are about pitfalls and opportunities in lie detection. In Chapter 15 I discuss the opportunities and elaborate on how people could improve their lie detection skills. In this chapter I focus on the pitfalls and outline why people fail to catch liars (see also Vrij, 2004a,b, 2007). Although most of the 15 reasons why people fail to detect deceit have already been discussed throughout this book, in this chapter I present them in a more systematic manner. To reduce repetition, I discuss each briefly and refer to the relevant chapters for more information. I believe that the reasons why lies remain undetected can be clustered into the following three different categories (Vrij, 2007): (i) a lack of motivation to detect lies; (ii) difficulties associated with lie detection; and (iii) common errors made by lie detectors.

LACK OF MOTIVATION TO CATCH LIARS

(1) The Ostrich Effect

Lies often remain undetected because people do not attempt to find out the truth (Chapter 1). I labelled this the *ostrich effect*. People are sometimes not motivated to learn the truth, because it is not in their best interest to do so. People generally appreciate the compliments they receive about their body shape, their hairstyle, their achievements, the way they dress, and so on. So why bother trying to discover whether those compliments are spoken in truth?

More serious lies also remain undiscovered for this reason. I mentioned in Chapter 1 that Betty Currie, who was Bill Clinton's secretary at the time of his presidency, tried to avoid learning details of the relationship between the President and Monica Lewinsky. Indeed, rather than gain anything from knowing the truth, it would have put her in the difficult position of having to decide what to do with such knowledge. Not knowing what to do when having learnt the truth may also be the reason why people often do not try to discover possible infidelity in their romantic partners. Discovering that their partner is having an affair will most likely cause a difficult situation for the betrayed spouse, and there is a risk, when confronted with the evidence, that the cheating partner will decide to leave, and the betrayed spouse may think that a split up is not in the interest of their children.

LIE DETECTION IS A DIFFICULT TASK

The second cluster of reasons as to why people fail to catch liars is that lie detection is a difficult task. I discuss seven reasons belonging to this category.

(2) Absence of Pinocchio's Growing Nose

Pinocchio's nose was unaltered each time he spoke the truth, but grew larger each time he lied. As such, his growing nose was a reliable cue to deceit. Throughout this book I hope it has become clear to the reader that there are no behavioural, speech, or physiological cues uniquely related to deceit. In other words, reliable cues to deception akin to Pinocchio's growing nose do not exist. This makes lie detection difficult because there is no cue that the lie detector can truly rely upon.

This does not mean that cues to deception never occur. They do occur, but different people may display different cues to deceit, and the same person may display different cues in different situations. In addition, the research literature discussed throughout this book has revealed that some cues are more likely to occur when people lie than others. For example, it is somewhat more likely that people will refrain from making subtle movements with their hands and fingers when they lie than that they will avert their gaze (Chapter 3); and some areas of the brain are more likely to be activated when people lie than other areas of the brain (Chapter 13).

(3) Subtle Differences

Another difficulty that lie detectors face is that if differences between truth tellers and liars occur they are typically very small. For example, facial expressions of emotions may sometimes betray a lie but such expressions can be as brief as 1/25th of a second (Chapter 3). Differences in movements displayed by truth tellers and liars are often subtle too (Chapter 3), and sometimes a lie is given away by only a tiny slip of the tongue.

In that respect Freud's (1959, p. 94) quote that "No mortal can keep a secret. If his lips are silent, he chatters with his finger-tips; betrayal oozes out of him at every pore" is misleading. I find that police manuals are also misleading when they describe nonverbal and verbal cues to deceit. Although such manuals typically contain small warnings about the unreliability of cues to deception, this is easily lost in the subsequent detailed and enthusiastic descriptions of how behaviour and speech differs between truth tellers and liars (see also Moston, 1992). Those descriptions are sometimes accompanied with pictures that have headings such as "truthful forward posture" and "deceptive adaptor behaviours" (Inbau, Reid, Buckley, & Jayne, 2001, pp. 145 & 149), which all give the suggestion that (i) reliable cues to deception exist and (ii) the differences between truth tellers and liars are substantial and therefore easy to spot.

(4) Countermeasures

A further complication for lie detectors is that liars sometimes deliberately attempt to appear credible to avoid detection. Methods used to achieve this are called *countermeasures*. I discussed that liars can avoid being detected in a polygraph examination by employing easy to learn countermeasures, such as pressing the toes against the shoe sole, biting the tongue, or thinking about a frightening event (Chapters 11 and 12). If liars use the appropriate countermeasures, they can also avoid detection when their brain activities are measured via EEGs or brain scans (Chapters 12 and 13); and people who are informed about the working of the verbal veracity assessment tool CBCA can tell stories that sound plausible to CBCA experts (Chapter 8).

I have not yet discussed whether people can successfully employ countermeasures to conceal nonverbal cues to deception, and I am aware of only two experiments that address this issue (Caso, Vrij, Mann, & DeLeo, 2006; Vrij, Semin, & Bull, 1996). In both studies, some participants were informed that liars are inclined to decrease their hand and finger movements (this is one of the more reliable findings that have emerged from the nonverbal deception literature, Chapter 3), whereas others received no information. Participants subsequently told the truth or lied. This information had no effect on the number of hand and finger movements made by the participants, and informed liars still made fewer of such movements than truth tellers. These findings suggest that nonverbal countermeasures may be not as easy to apply as verbal and physiological countermeasures, but more research is needed to demonstrate this.

Obviously, the successful use of countermeasures is a major challenge for professional lie detectors. If it is known amongst terrorists, spies and criminals which lie detection tools will be used to catch them, they may learn more about these tools and attempt to beat them. If they succeed in doing so, the tools are no longer effective.

(5) Embedded Lies

Another difficulty that lie detectors face is that lies are often embedded in truths (Chapters 8 and 9). That is, rather than telling a blatant lie that is entirely beyond the truth, people tend more to change some vital details in an otherwise truthful story. Thus, when a man wants to cover up his activities on Tuesday night, he could describe what he did on another night (say Monday night) but pretending that those events happened on the Tuesday night in question. Therefore, if he went to the gym on Monday night he could mention that he went to the gym on Tuesday night, and subsequently describe his experiences at the gym on Monday night. Most of the statement is then truthful, with only a tiny, but vital, lie embedded. My impression of watching videotapes or listening to audiotapes of police–suspect interviews is that suspects often tell such embedded lies (see also Hartwig, Granhag, & Strömwall, 2007; Strömwall, Granhag, & Landström, 2007). In a similar vein, when examining false identities given by criminals, Wang, Chen, and Atabakhsh (2004) found that criminals typically change only a small portion of their original identity.

I expect that when people lie outside a criminal context they use a similar embedded lies strategy (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). This strategy makes lying easier, because the liar does not need to fabricate a plausible story or remember many fabricated details, but it makes lie detection more difficult. It hampers lie detectors who examine verbal statements because they typically examine the quality of detail included in a statement (Chapters 8 and 9). Lies that are embedded in predominantly truthful statements may contain many high quality details, which could give the lie detector the erroneous impression that the statement is truthful. Lie detectors who examine nonverbal behaviour may make a similar mistake if the deceptive part of a liar's story remains unnoticed (e.g., *when* the person went to the gym), and if they concentrate on the truthful part instead (e.g., what the person did at the gym).

Lie detectors who examine physiological responses are not affected by embedded lies because their examinees often do not have the opportunity to tell embedded lies. Instead, they answer specific questions such as "Were you at the gym last Tuesday night between 8 and 9 pm?"

(6) No Adequate Feedback

Another complication in lie detection is that lie detectors often do not receive adequate feedback about their judgements and therefore cannot learn from their mistakes (Chapter 5). For feedback to be adequate it needs to be provided frequently, reliably, and immediately. Thus, observers should be informed immediately after every interaction with another person whether that person was lying or not. They could then work out how liars truly behave, what they really say, and, if measured, how they respond physiologically. However, adequate feedback is often not available. People often do not discover whether or not they have been lied to, and in case they do come to know this, it is often a long time after the interaction took place (Chapter 6). By the time they learn that a person has lied to them, they have probably already forgotten how the person responded exactly. Criminal investigators could try to address this problem by videotaping the people they talk to or by keeping charts of examinees' physiological responses. When they come to know the truth at a later stage, they then could go back to this material to check the interviewee's responses. I am not aware of this happening in real life. In addition, this feedback is unlikely to be adequate because it may be biased. As I explained in Chapter 11, polygraph examiners often do not come to know the incorrect decisions they made. Polygraph examinations are carried out in cases where no evidence is available, and therefore a confession is used to verify the examiner's decisions. An examinee is considered guilty if he or she confesses to the crime, and is considered innocent if someone else confesses to the crime. However, a guilty examinee who passes the polygraph test (false-negative error) is unlikely to confess, because there is no evidence against him or her. It is unlikely that someone else will confess to this crime and this means that the case probably remains unresolved, and so the examiner will not learn of his error. An innocent person who fails the polygraph test (false-positive error) may falsely confess if he or she is submitted to lengthy interrogations after failing the polygraph test. In that case the feedback the examiner receives is incorrect and the error will not be discovered. Alternatively, the innocent examinee may deny involvement in the crime, but since the police will think that the examinee is the guilty person, they are unlikely to arrest and interrogate someone else. The case is therefore likely to remain unresolved and the examiner will not come to know the mistake either.

(7) Violation of Conversation Rules

Lying will become increasingly difficult if the lie detector persists in asking questions. When answering these questions, liars must make sure that they sound plausible and must avoid contradicting themselves; should not say things the observer knows or may find out to be untrue; must remember what they have said so that they are consistent when re-telling their story; and must avoid making slips of the tongue or providing new leads. In addition, they must monitor their own behaviour during the conversation so that they do not give away their lies via non-verbal cues, whilst monitoring the reactions of the lie detector so that they can adapt their response strategy if their current strategy raises suspicion (Chapter 3). However, in daily life conversations it is often seen as inappropriate, strange or impolite if the listener asks further questions. Conversation partners will probably not appreciate utterances such as "Could you elaborate on that?", "Could you back this up with evidence?", "Could you repeat what you just said?" and may object to it.

Also, it would benefit the lie detector if he or she examined the speaker's movements because they may reveal signs of deceit, but this would be considered inappropriate in daily life situations. Conversation rules dictate that a listener looks the speaker into the eyes, yet the eyes generally do not reveal reliable information about deception (Chapter 3). Therefore, the two conversational rules of (i) not asking further questions and (ii) looking someone into the eyes hamper lie detection and make it a difficult task.

(8) Good Liars

The final reason that makes lie detection difficult is that some people are very good liars. Surprisingly, there is little research regarding what makes someone a good liar, but I believe that three criteria are relevant (see also Vrij, Granhag, & Mann, in press). Good liars are those people: (i) whose natural behaviour disarms suspicion; (ii) who do not find it cognitively difficult to lie; and (iii) who do not experience emotions such as fear, guilt, or duping delight when they are lying (Chapters 3, 5 and 6). I think that these three criteria include eight characteristics: (i) being a natural performer; (ii) being well prepared; (iii) being original; (iv) thinking rapidly; (v) being eloquent; (vi) having a good memory; (vii) not experiencing feelings of guilt, fear, or duping delight while lying; and (viii) being good at acting.

Natural performers

Certain behavioural patterns are associated with honesty and likeability, such as directed gaze to a conversation partner, smiling, head nodding, leaning forward, direct body orientation, posture mirroring, uncrossed arms, articulate gesturing, moderate speaking rates, a lack of ums and ers, and vocal variety (Buller & Aune, 1988; Ekman, 1985/2001; Tickle-Degnen & Rosenthal, 1990). Some people show such demeanour naturally even when they are lying (e.g., *natural performers*, Ekman, 1997). These natural performers are likely to be good liars because their natural behaviour is likely to allay suspicion.

Being well prepared

Good liars are probably *well prepared* and have worked out in advance what they are going to say and do. The more thorough their preparation, the easier lying probably becomes. Obviously, liars should prepare a lie that sounds plausible. Elsewhere we described five cases of people who were suspected of having killed one of their relatives and initially denied having done so (Vrij & Mann, 2001b). Some of those people made serious mistakes when they planned their stories which made it easy to find out that they probably were hiding the truth. For example, in one case blood stains suggested that the victim was actually killed somewhere other than where claimed. Another person reported to have been knocked unconscious for 10 hours, but anaesthetists said that this was impossible.

As these examples demonstrate, lie detectors are likely to check a person's story and search for evidence that confirms or contradicts it. Good liars therefore *say as little as possible* or *say things that are impossible for others to verify*. The less verifiable information is given, the less opportunity it provides for the lie detector to check.

Being original

Liars can always face unexpected situations that require an explanation. For example, a wife may confront her husband with the telephone number and address of a woman – unknown to her – which she found in his pocket; or the police detective may tell the suspect that he was seen by a witness at the scene of crime directly after the crime took place. To lie successfully in these or similar situations, a convincing and plausible answer is needed. To spontaneously invent a plausible answer is probably too difficult for many liars, but *original thinkers* who are *mentally creative* may be able to succeed.

Rapid thinking

The answer a liar gives to an unexpected question may sound plausible, but it will still raise suspicion if the liar waits too long before giving that answer. The ability to *think rapidly* is thus another skill that good liars should possess.

Being eloquent

Good liars are also *eloquent* speakers. People who typically use many words to express themselves may find it easier to talk themselves out of an awkward situation. They can commence with giving a circumlocutionary response, which in fact does not answer the question, whilst thinking about the appropriate answer. Or they can use their eloquence to fool the listener, by giving a response which sounds convincing but which, in fact, does not provide an answer to the question. Many politicians are very good at this.

Good memory

Liars always run the risk of lie detectors asking them to repeat or clarify what they have said before. They then need to remember what they said before so that they can repeat the same story or add information without contradicting their previous statement. A *good memory* is thus required. Liars could make the memory task easier for themselves by *saying as little as possible*. The less someone says, the less that person has to remember. In addition, the memory task will become easier if the liar *stays with the truth as much as possible*. In other words, rather than fabricating a whole story, a liar could tell a story which is mostly true but differs from the truth in only a few crucial details (see the embedded lies section).

Not experiencing guilt, fear or delight

Liars differ in the emotions they experience. One applicant may feel guilty when exaggerating her current salary, whereas another applicant may not experience any guilt while doing this. One suspect may experience detection anxiety when presenting a fake alibi, whereas another suspect will remain calm. One pupil may be delighted when realising that the headteacher is about to believe his excuse for being late, whereas another does not experience such duping delight. Deceiving others is made easier if the liar does not experience feelings of guilt, fear or delight, because in that case there will not be any emotional behaviour that needs to be suppressed. An absence of emotions during deception is possibly related to *being practiced at lying* and *feeling confident when lying*. Moreover, people are unlikely to experience guilt or fear if they have a strong imagination and therefore the capacity to *believe what they are saying.*

Good at acting

Although natural performers and those who do not experience cognitive load or emotions when lying make the best liars, those who could properly mask signs of cognitive load and emotions and, at the same time, display behaviour that appears credible, probably make good liars too. This requires *good acting skills*. These good actors should also have good *decoding skills*. Since they are not natural performers, their lies may raise suspicion and they should adapt themselves adequately to disarm this suspicion. The sooner they adapt themselves the more chance they have of successfully disarming suspicion. It is thus crucial to notice suspicion quickly which requires good decoding skills.

COMMON ERRORS MADE BY LIE DETECTORS

People do not only fail to catch liars because they are unmotivated to catch them or because the lie detection task is difficult, they also make errors. Different people make different errors and therefore many errors can be identified. I believe that these errors can be divided into seven groups, each of which I shall now examine in this section.

(9) The Othello Error

A common error in lie detection is to interpret some signs too easily as cues to deception. This happens in particular with signs of nervousness. The mistake that lie detectors then make is not considering that truth tellers can be nervous too and may show nervous responses (Chapter 3). Such a wrong decision, mistakenly interpreting signs of nervousness displayed by truth tellers as signs of deceit, is called the *Othello error* (Ekman, 1985/2001), named after Othello, the title character in

Shakespeare's play. Othello falsely accuses Desdemona (his wife) of infidelity. He tells her to confess since he is going to kill her for her treachery. Desdemona asks Cassio (her alleged lover) to be called so that he can testify her innocence. Othello tells her that he has already murdered Cassio. Realising that she cannot prove her innocence, Desdemona reacts with an emotional outburst which Othello misinterprets as a sign of her infidelity.

Police manuals often advise their readers to pay attention to cues of nervousness when attempting to detect deceit (Chapter 5), and following this advice could easily lead to Othello errors. Also in the Behaviour Analysis Interview lie detectors are encouraged to look for signs of nervousness (Chapter 7), increasing the likelihood of making Othello errors. Concern-based polygraph tests are also vulnerable to the Othello error. As I explained in Chapter 11, it is not unreasonable to believe that innocent suspects will be more concerned when answering the relevant questions (e.g., "On 26 January, did you shoot Julie Pearchap?") than when answering the probable lie questions (e.g., "Have you ever physically hurt someone in your life to get revenge?"), particularly because polygraph tests are conducted when no other evidence is available. This means for innocent suspects that they cannot prove their innocence. The stakes are thus high for them when answering the relevant questions, and they may well be more concerned when answering them than when answering the probable lie questions.

(10) Examining the Wrong Cues

There are widespread, but often incorrect, beliefs about how people respond when they lie (Chapter 5). Both laypersons and professional lie catchers overwhelmingly expect liars to react nervously, with "liars look away" and "liars make grooming gestures" amongst the most popular beliefs. Such cues are not reliable cues to deception (Chapter 3). I gave several explanations from where such beliefs originate (Chapter 5), and one of those explanations was that people are taught such incorrect cues. For example, in their influential police manual Inbau *et al.* (2001) promote the idea that liars look away and make grooming gestures. In fact, they discuss around 10 nonverbal cues as being diagnostic cues for deception, whereas only one has been found to be associated with deception according to research (Chapter 3): lying is associated with a decrease in illustrators. It is therefore not surprising that in our lie detection study where police officers saw video fragments of suspects who told the truth or lied during their police interviews, we found that the more the police officers endorsed the lie cues promoted in the Inbau *et al.* manual, the worse they became at distinguishing between suspects' truths and lies (Chapter 6).

(11) Neglect of Interpersonal Differences

There are large individual differences in people's behaviour, speech, and physiological responses. Some people typically make many movements, others do not; some people are eloquent, others are not; some people show large variations in physiological responses, others do not, etc. Verbal lie detection tools such as SVA attempt to control for these interpersonal differences by using a Validity Checklist (Chapter 8), and, with the exception of thermal imaging, physiological lie detection tools also attempt to control for interpersonal differences by asking comparison questions (Chapters 11 to 13).

However, when evaluating behavioural responses, a control for interpersonal differences often does not take place (Chapters 3 and 6). The result is that people whose natural behaviour looks suspicious (e.g., people who naturally avert their gaze or fidget a lot) are in a disadvantageous position, because they run the risk of being falsely accused of lying. In Chapter 6, I explained that introverted and socially anxious people in particular run such a risk. Errors are also easily made when people of different ethnic backgrounds or cultures interact with each other, because behaviours naturally displayed by members of one ethnic group or culture may appear suspicious to members of another ethnic group or culture (Chapter 6).

(12) Neglect of Intrapersonal Differences

Not only do different people respond differently in the same situation (*inter*personal differences), the same person also responds differently in different situations (*intra*personal differences). Neglecting or underestimating those intrapersonal differences is another error that lie catchers make. The failure to control adequately for intrapersonal differences is one of the main criticisms of concern-based polygraph tests (Chapter 11). As I explained above in the Othello-error section, the comparison question: "Have you ever physically hurt someone in your life to get revenge?" is not necessarily an adequate control for the relevant question: "On 26 January, did you shoot Julie Pearchap?" Moreover, in police interviews police officers are advised to examine a suspect's natural, truthful, behaviour during the small-talk part at the beginning of the interview and to compare this behaviour with the behaviour shown by the suspect during the actual interview. Differences in behaviour could then be interpreted as "significant" (Inbau *et al.*, 2001). Although the idea sounds appealing, this approach is prone to incorrect judgements because an incongruent comparison is made. Small-talk and the actual investigation are fundamentally different situations. Small-talk conversations are low-stakes situations where the suspect's responses are unlikely to have any negative consequences. In contrast, the investigative part of the interview is a high-stakes situation where the suspect's reactions do matter. If the suspect's reactions raise suspicion, he or she may be suspected of having committed the crime which could lead to serious negative consequences. Therefore, unsurprisingly, both guilty and innocent suspects tend to show different behaviours during small-talk compared to the actual interview (Vrij, 1995).

The tendency to neglect or underestimate the importance of intrapersonal differences is not an error made only by lie detectors. It is a well-known error in social perception and called the *fundamental attribution error*, the tendency to overlook the impact of situations when explaining someone's responses (Ross, 1977).

(13) The Use of Heuristics

Rather than carefully scrutinising someone's responses when attempting to detect deceit, observers may instead rely on general decision rules. Person perception researchers have emphasised that this is the most effective way for observers with limited time and attentional resources to deal with complex environments (Macrae & Bodenhausen, 2001). However, general decision rules, often labelled *heuristics* by deception researchers, easily lead to systematic errors and biases. Several heuristics thought to influence veracity judgements can be distinguished, and I discussed many of them in Chapter 6 (see Cole, Leets, & Bradac, 2002; Fiedler & Walka, 1993; Levine, Park & McCornack, 1999, and Vrij, 2004b, for reviews).

In daily life, people encounter more truthful than deceptive messages, and they therefore assume that the behaviour they observe is usually honest (the *availability heuristic*, O'Sullivan, Ekman, & Friesen, 1998). Related to this is the *anchoring heuristic* (Elaad, 2003). It refers to people making insufficient adjustments from an initial value (the anchor) resulting in a final decision that is biased towards this value. Thus, if observers are pre-occupied in thinking that someone is telling the truth, they will make insufficient adjustments when contrasting evidence emerges. It has further been argued that as romantic relationships become more intimate, partners develop a strong tendency to judge the other as truthful, the so-called *relational truth-bias heuristic* (Stiff, Kim, & Ramesh, 1992). The *probing heuristic* (Levine *et al.*, 1999) refers to observers' tendency to believe a source more after the source has been probed. (Observers believe that probing is an effective lie detection strategy; if probing does not result in clear signs of deceit, and it often will not, the source is more likely to be believed.) The *representativeness heuristic* (Stiff, Miller, Sleight, Mongeau, Garlick, & Rogan, 1989) refers to the tendency to evaluate a particular reaction as an example of a broader category. Used in a deception context it could explain people's inclination to interpret nervous behaviors as signs of deceit. The *expectancy violation heuristic* (Vrij, 2004b) refers to the tendency to judge reactions which are odd or infrequent (e.g., keeping the eyes closed, or conversely, staring during a conversation) as deceptive. According to the *falsifiability heuristic*, messages that contain more falsifiable details are more likely to be judged deceptive (Fiedler & Walka, 1993). The *facial appearance heuristic* (Vrij, 2004b) refers to the tendency to judge people with attractive faces or baby-faced appearances as honest.

O'Sullivan (2003) demonstrated that the *fundamental attribution error* (introduced above) undermines lie detection. When people form impressions of others, they tend to overestimate dispositional (e.g., "character") factors of that person and tend to underestimate situational factors. Therefore, when an observer believes that someone is generally a trustworthy person, he or she will tend to judge that person as truthful in any given situation. In contrast, when the observer believes that someone is an untrustworthy person, he or she will be inclined to judge that person as dishonest in any given situation. Obviously, trustworthy people are not honest all of the time and untrustworthy people are not always dishonest.

(14) Poor Interview Techniques

Several interview techniques promoted in police manuals hamper lie detection. For example, police detectives are sometimes advised to confront suspects at the beginning of the interview with the evidence they have gathered earlier on in the investigation process (Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Hartwig, Granhag, Strömwall, & Vrij, 2005; Leo, 1996a). This tactic is designed to show suspects that it is fruitless to remain silent and that they are better off confessing. This interview style will hamper lie detection. One of the difficulties liars face is that they do not know what the observer knows. They therefore do not know what they can say without running the risk of contradicting facts known to the observer. If police officers disclose the facts they know to suspects, they reduce the uncertainty for lying suspects and make the lying process a more comfortable experience. Disclosing evidence early on provides liars with the opportunity to adapt their stories and to give an innocent explanation for the evidence. Thus, if a culprit is told that CCTV footage showed that she was seen at Commercial Road on Saturday afternoon at 3pm, she can now construct an alibi that gives an innocent explanation for her presence there on that Saturday afternoon.

Another unfortunate strategy from a lie detection perspective is if police detectives accuse suspects of lying. This gives suspects who actually lie the ideal opportunity to "escape" from the interview situation by saying that they will no longer cooperate with the investigation, claiming that further cooperation is futile because they are not believed anyway. Also, accusing someone of lying may elicit the same responses in liars and truth tellers. That is, suspects correctly accused of lying may experience detection anxiety, however, suspects incorrectly accused of lying may also become afraid of not being believed (Ofshe & Leo, 1997). Because of that fear, truth tellers may show the same nervous behaviours as liars (Bond & Fahey, 1987). This puts the lie detector who accuses the interviewee in a difficult position: should signs of fear be interpreted as signs of guilt or as signs of innocence? The behaviour does not provide the answer (see also the Othello error above).

(15) Overestimating the Accuracy of Lie Detection Tools

The final error that I want to point out is that professional lie catchers tend to overestimate the accuracy of the lie detection tools they use. This book has made clear that every single lie detection tool used to date is far from accurate and prone to errors. Despite the fallibility of the tests, users claim high accuracy. I gave several examples in Chapter 1 and throughout this book. Despite the difficulty in judging veracity on the basis of someone's nonverbal and verbal behaviour, and despite the sometimes simplistic techniques used to judge veracity on the basis of demeanour, practitioners claim 85% accuracy by using these techniques (Chapters 1 and 7). Another technique is thermal imaging whereby people's arousal is measured via the blood flow around the eyes. The technique does not take interpersonal or intrapersonal differences into account and is therefore prone to many errors. Yet, it has been suggested that the technique could be used at airports as a possible screening device (Chapters 1 and 11). Concern-based polygraph tests are error prone due to the difficulty of asking comparison questions that are truly comparable to the relevant questions. Yet, when I talk to polygraph examiners they typically claim 95% accuracy of the test (see also Chapter 1). Using fMRI brain scans is the latest technology to join the lie-detecting arsenal. After testing the technique with relatively few undergraduate students under ideal testing circumstances (low-stakes scenarios in which comparable questions are easy to formulate, and testing without delays so that the guilty examinees clearly remember the details they are instructed to lie about) developers of the tool claim that the test is now "ready to detect terrorists" (Chapters 1 and 13).

The danger is that if these practitioners believe their own claims, and there is no reason to believe that they do not, they will try to convince others that their techniques work and can be relied upon. These others could include people who are perhaps not familiar with all the details about lie detection but who have great interest in catching liars, such as civil servants working for Ministries of Defence or National Security Departments, airport directors, judges and so on. If they are convinced that the tests work, they may introduce them and uncritically rely on them, possibly resulting in terrorist attacks, miscarriages of justice, and other undesirable outcomes.

CONCLUSION

In this chapter I presented 15 reasons as to why people fail to catch liars. One reason was that people are sometimes not interested in coming to know the truth; other reasons were related to the difficult nature of a lie detection task, and to errors lie detectors commonly make. Being unaware of the truth suits people well in many daily life interactions (Chapters 2 and 6), but not in all. I believe that people would become better in discriminating between truths and lies if they were more aware of the difficulties they face and hence were to avoid making errors. In the next chapter I provide guidelines on how to do this.

CHAPTER 15

Opportunities: How People Can Improve Their Lie Detection Skills

In this final chapter I discuss three topics. First, I briefly summarise how accurate lie detectors are at distinguishing between truths and lies when they use the various lie detection tools presented in this book. I will not discuss these tools again here, as I have already done so in Chapters 6 to 13. This section shows that, at least in laboratory studies, lie detectors perform worse when they observe someone's nonverbal and verbal behaviour, and perform best when they measure physiological responses; detailed analyses of speech fall in between those two approaches.[1] Second, I make some suggestions about how to improve lie detection via analysing speech or measuring physiological responses. I will argue that collaboration is needed between deception researchers and researchers in other areas of psychology.

Third, most of this chapter deals with how to improve lie detection via observing nonverbal and verbal behaviour, and I present 17 guidelines that could be used to detect deceit in this manner. In this chapter I pay more attention to this type of lie detection than to lie detection

[1]When the speech content is systematically analysed via verbal veracity assessment tools I refer to *speech analysis*; unsystematic observations of speech are referred to as observing someone's *verbal behaviour*.

based on speech analysis or physiological responses and I do this for two reasons: (i) it is the most inaccurate type of lie detection, and therefore perhaps mostly in need of improvement; and (ii) this form of lie detection can be used in many more situations than the other two types of lie detection, because it does not require equipment such as a polygraph, EEG recording electrodes, fMRI brain scan (necessary for physiological lie detection), or transcribing someone's speech (necessary for speech analysis). As I did in Chapter 14, I will simply refer to previous chapters where appropriate to avoid unnecessary repetition.

ACCURACY OF THE DIFFERENT LIE DETECTION METHODS

Throughout this book I discussed the accuracy rates obtained in both laboratory and field research. In laboratory research observers detect truths and lies told by (typically) undergraduate students for the sake of the experiment. In field studies observers detect truths and lies told in real-life settings. Table 15.1 presents a summary of the accuracy rates in laboratory and field studies. Field studies that report on ability to detect truths and lies via observing verbal and nonverbal behaviour or via speech analysis are rare, and the outcomes of these few studies are often unreliable. I therefore do not discuss them here. Field studies in physiological lie detection are more frequent, particularly regarding CQT polygraph tests, and I will present a summary of the outcomes in this section. However, as I explained in Chapter 11, caution is necessary when interpreting these outcomes. In many of these field studies the ground truth could not be established with certainty, and a further problem was the sampling bias. The consequence of the sampling bias is that the accuracy scores in polygraph field studies may well be higher than would actually be found in real life.

Laboratory Studies

Table 15.1 shows that people perform the worst when they attempt to distinguish truths and lies from somebody's nonverbal and verbal behaviour. The average accuracy scores found in laboratory studies with professional lie catchers (56%) are only just above what could be expected by flipping a coin (Chapter 6). Judging veracity by analysing written transcripts of speech results in accuracy scores of around 70%, with not much difference between using Criteria-Based Content Analysis (CBCA, Chapter 8) or Reality Monitoring (RM, Chapter 9).

Table 15.1 Accuracy rates in truth and lie detection of various lie detection tools[1]

	Accuracy rate (%)[2]		
	Truth	Lie	Total
Nonverbal and verbal behaviour			
Laypersons (laboratory studies)[3]	63	48	
Professionals (laboratory studies)[4]	56	56	
Speech analysis			
CBCA (laboratory studies)	71	71	
RM (laboratory studies)	72	66	
Physiological analysis			
CQT polygraph (laboratory studies)	60–84	74–82	
GKT polygraph (laboratory studies)	83–99	76–88	
EEG-P300 utilizing the GKT (laboratory studies)	88	82	
fMRI			78–93[5]
CQT polygraph (field studies)	53–75	83–89	
GKT polygraph (field studies)	94–98	42–76	

Notes: [1]Insufficient laboratory research has been carried out with BAI and SCAN in order to estimate their accuracy scores; [2]See for full details about the accuracy scores Appendix 6.1 (nonverbal and verbal behaviour, laypersons), Table 6.3 (nonverbal and verbal behaviour, professionals), Table 8.3 (CBCA), Table 9.2 (RM), Table 11.2 (CQT polygraph laboratory studies), Table 11.3 (CQT polygraph field studies), Table 12.2 (GKT polygraph laboratory studies), Table 12.3 (GKT polygraph field studies) and Table 12.5 (EEG-P300 using a GKT method). The fMRI accuracy scores are reported in the main text of Chapter 13; [3]Laypersons detecting truths and lies told by adults they do not know; [4]Professionals detecting truths and lies told by adults they do not know; [5]For fMRI scores only total accuracy scores rather than truth and lie accuracy scores have been reported

In Chapters 11 and 12 I introduced two main types of questioning techniques in physiological lie detection: the Comparison Question Test (CQT) and the Guilty Knowledge Test (GKT). Traditionally, during those tests examinees' electrodermal activity (EDA), respiration and blood pressure are measured with a polygraph. I presented in Chapters 11 and 12 the laboratory research reviews published to date examining the accuracy of CQT and GKT polygraph testing. Since different researchers included different studies in their reviews, the accuracy rates differed, and I have made this evident in Table 15.1 by reporting the lowest and highest accuracy rates obtained in those reviews. Despite the differences between the reviews, it becomes clear that the accuracy rates obtained with both types of polygraph testing are typically higher than those achieved by using CBCA and RM. Recently, rather than examining EDA, respiration and blood pressure, researchers started to measure examinees' P300 brain waves (with EEGs) or brain activity in different areas and structures of the brain (with fMRI scans). The accuracy rates for EEG-P300 and fMRI tests in laboratory studies are similar to those found in laboratory polygraph testing.

Field Studies

Sometimes very high accuracy rates are reported in laboratory-based physiological lie detection studies, and this has led to strong claims such as, "these tests are ready to detect terrorists" (Chapters 1 and 13). As I explained in Chapters 11 to 13, those laboratory studies are typically carried out under optimum testing conditions which mask the problems associated with CQT and GKT examinations. A major problem with the CQT is to design comparison questions that are true controls for the relevant questions; a main problem with the GKT is to select critical details that the guilty examinee is likely to remember. Masking problems implies that in real life the accuracy rates are probably lower than in laboratory studies.

The polygraph field study accuracy scores presented in Table 15.1 show that this is the case, particularly when considering that in real life the accuracy rates may be lower than in those field studies due to the sampling bias (Chapter 11). The accuracy rates reflect the problems associated with the CQT and GKT. The CQT is vulnerable to false-positive errors, that is, falsely accusing truthful examinees of lying. This occurs when the comparison questions are not an adequate control for the relevant questions (Chapter 11). Difficulty in pinpointing truth tellers is the result and in no review of field studies have truth accuracy rates higher than 75% been reported. The GKT is vulnerable to false negative errors, that is, classifying guilty examinees as innocent. This happens when guilty examinees do not recognise the critical details (Chapter 12). Difficulty in pinpointing liars is the result and in neither of the two GKT polygraph field studies published to date have lie accuracy rates higher than 76% been obtained. The field studies have only been carried out in polygraph examinations. There are no theoretical reasons to assume that the accuracy rates of EEG-P300 and fMRI examinations will be higher in real life than those of polygraph examinations.

Seriousness of False-Positive and False-Negative Errors

Both false-positive and false-negative errors are undesirable but which is the most serious depends on the situation. In criminal investigations (catching thieves, rapists, murderers, etc.) false-positive errors are considered as the most serious errors (Chapter 11) because they could lead to the prosecution and conviction of innocent people. In contrast, one could argue that in security threat investigations (catching spies or terrorists) false-negative errors are the most serious errors because making such errors would mean that spies and terrorists remain undetected. However, as I made clear in Chapter 11, for lie detection tools to be useful in catching spies and terrorists they need to be highly accurate, and far more accurate than they currently are. A high accuracy rate is needed because of the low base rate of potential spies and terrorists in a given sample. Suppose that there are 10 spies amongst 1000 employees, and further suppose that the lie detection test has 90% accuracy. In theory, it means that most but not all spies (nine) will fail the test but also that a further 99 non-spies will fail the test (10% of the 990 non-spies). The lie detector thus ends up with 108 persons who failed the test, and no further opportunity to discriminate between the nine spies and 99 non-spies on the basis of the test outcomes.

In summary, although physiological lie detection tools often result in higher accuracy rates than non-physiological lie detection tools (at least in laboratory studies), errors are frequently made with all types of test. This makes them all unsuitable to rely upon to unmask spies and terrorists. However, some tests can discriminate truths and lies above the level of chance and are therefore useful to form impressions about someone's veracity in criminal investigations, as I outlined in Chapters 6 to 13.

SUGGESTIONS FOR IMPROVEMENT IN SPEECH ANALYSIS AND PHYSIOLOGICAL LIE DETECTION

I think that progress in speech analysis and physiological lie detection can be made in four areas in particular. First, I think that in both types of lie detection properly conducted field studies are urgently needed. Only after carrying out field studies will we know how accurate such lie detection tests really are, and what their true strengths and weaknesses are. I appreciate the difficulties in conducting field studies, but without such externally valid studies we have little idea of how effective deception detection tools that are currently in use actually are. Lie detection tools such as SVA (Chapter 8) are used as evidence in criminal courts while nobody knows how accurate SVA assessments are. Huge numbers of criminal investigators are trained to carry out BAI interviews (Chapter 7), and the SCAN method is taught worldwide (Chapter 10), but it is unknown how accurate investigators who employ these techniques are. Finally, EEG-P300 and fMRI examinations

(Chapters 12 and 13) are presented by some researchers as *the* method to catch criminals and terrorists, but these lie detection tools have only been tested under optimum conditions in controlled and confined laboratory settings with undergraduate students as participants.

Second, regarding speech analysis lie detection, it may be useful to combine elements of the CBCA and RM tools as I indicated in Chapter 9. I also believe that the Validity Checklist, used by verbal lie detectors to assess the CBCA outcome, needs further development. As I explained in Chapter 8, the Validity Checklist contains several criteria that are in my view not well defined or justified, whereas other issues that justify inclusion in the Validity Checklist are not included.

Third, regarding physiological lie detection, the CQT is not a standardised test (Chapter 11). Instead, for each examinee the examiner designs a specific test. This procedure could be seen as a very sophisticated piece of psychological engineering (Lykken, 1998). A range of factors can influence the mindset of both examiner and examinee which could subsequently determine the examinee's physiological responses and thus the outcomes of the test. Also the wording of the questions and the order in which they are asked could influence the outcomes (Chapter 11). The lack of standardisation of the CQT is not acknowledged enough by CQT researchers and examiners. I believe that a greater understanding of social influence and linguistics research would help them identify factors that could have an impact on test outcomes.

For a GKT to work, guilty examinees must recognise the critical details presented to them during the test (Chapter 12). The problem that GKT examiners face is that they cannot know which critical details guilty examinees are likely to remember and thus will recognise during the test. Eyewitness testimony theory and research have provided valuable insight into what people are likely to remember about events (although most of this research examines how *witnesses* rather than *offenders* perceive and remember crimes, Chapter 12). A greater understanding of the eyewitness literature could benefit designing a GKT.

In summary, I believe that both the CQT and GKT questioning techniques could be enhanced if research findings from disciplines outside the area of lie detection were incorporated in the techniques. The problem for CQT and GKT examiners and researchers is that there are so few researchers active in this area. A close examination of the researchers that I cited in Chapters 11 to 13 reveals that the area of physiological lie detection appears to be led by just a few (groups of) researchers (see also Iacono, 2000, and the National Research Council, 2003). Those few researchers cannot possibly be aware of all the published research that may be relevant to their questioning techniques. I therefore encourage them to seek collaboration with social influence, linguistic and eyewitness testimony researchers.

Fourth and finally, I believe that fMRI research could enhance our theoretical understanding of deception and the detection of deception. fMRI researchers could strive to establish which brain structures and areas are associated with different types of lies. This is a laborious task because lies can be classified by so many dimensions (e.g., low- vs. high-stakes; spontaneous vs. rehearsed lies; outright lies vs. exaggerations vs. concealments, etc.) and all combinations of these dimensions are possible. (These three dimensions already result in 12 different types of lies.) Because different people may show different patterns of brain activity, different groups of individuals need to participate in such research.

If brain activity profiles for different types of lie could be established, at least two further research activities appear relevant. First, it could be examined whether, and if so how, these neurophysiological brain activities relate to cues that are more easily accessible to lie detectors, such as psychophysiological cues (EDA, respiration, and blood pressure), speech-related cues, and behavioural cues (Happel, 2005). Second, as I will explain in the section below, lie detectors could employ interview protocols aimed at enhancing differences between truth tellers and liars. Researchers could design for each type of lie an interview protocol that creates the largest differences in brain activity between truth telling and lying. If brain activity is associated with psychophysiological, speech-related or behavioural cues, the larger the differences in brain activity between truth tellers and liars, the more likely it is that psychophysiological, speech-related and behavioural cues will discriminate between them.

GUIDELINES TO CATCH LIARS BY EXAMINING NONVERBAL AND VERBAL BEHAVIOUR

This guidelines section is divided into two parts.[2] In the first "Observational Guidelines" part, I present nine guidelines that deal with observing and interpreting the possible nonverbal and verbal cues that truth tellers and liars display. These observational guidelines are then followed by eight guidelines about interviewing to detect deception. Although it is not always possible to interview someone, in many cases

[2] See also Granhag and Vrij (2005); Vrij (2004a,b, 2006b, 2007); Vrij, Fisher, Mann, and Leal (2006, in press); and Vrij and Granhag (2007).

it is and I believe that employing specific interview protocols could enhance discrimination between truths and lies by observing someone's nonverbal and verbal behaviour. In the second "Interviewing to Detect Deception" part of this section I will explain how.

Observational Guidelines

(1) Use flexible decision rules

The research discussed in this book reveals that not a single behavioural or verbal cue is uniquely related to deception. In other words, there is no giveaway clue like Pinocchio's growing nose. Instead, different people show different cues to deception in a given situation (e.g., interpersonal differences); the same person shows different cues to deception on different occasions (e.g., intrapersonal differences); and sometimes people do not show cues to deceit at all.

This means that is inappropriate to use fixed decision rules based on individual cues, such as "liars look away" when attempting to detect deceit. In fact, we know that people who look at individual nonverbal or verbal cues are typically poor lie detectors (Mann, Vrij, & Bull, 2004; Vrij & Mann, 2001a). Instead, it is preferable to make veracity assessments on the basis of multiple cues (Ekman, O'Sullivan, Friesen, & Scherer, 1991; Vrij, Akehurst, Soukara, & Bull, 2004a; Vrij, Edward, Roberts, & Bull, 2000; Vrij & Mann, 2004). For example, it is not enough to conclude that someone is lying if he or she decreases his or her movements during a conversation, but the likelihood that the person is lying increases when that person's speech rate slows, and when the speech content becomes vaguer and less detailed. However, even such clusters of cues do not fit all liars nor do they fit a particular liar in all situations. In other words, fixed decision rules that include multiple cues are not satisfactory either. Instead, someone should use flexible decision rules that include multiple cues. That is, on one occasion a combination of movements, speech rate, and type of detail may betray a lie; on another occasion a combination of smiles, pauses, voice pitch change, and a slip of the tongue may give the lie away, etc.

(2) Cues to deceit are most likely to occur when liars experience emotions or cognitive load, or attempt to control themselves

In this book I made clear that the act of lying in itself does not affect somebody's nonverbal behaviour or speech. For differences between truth tellers and liars to emerge, liars often need to experience feelings of guilt or fear, or should be excited about the possibility of fooling someone (Chapter 3). Also, finding it difficult to lie or attempting to make a convincing impression on lie targets may result in differences between truth tellers and liars (Chapter 3). For many lies told in daily life the consequences of being disbelieved are minor (Chapters 2 and 3). In those low-stakes situations liars typically do not experience these emotions, do not find it difficult to lie, and do not put much effort into attempting to appear credible. Lie detection is therefore a daunting task in low-stakes situations because truth tellers and liars are unlikely to show different behavioural or speech cues.

High-stakes situations are different (Chapter 3). In high-stakes situations being disbelieved may have serious repercussions for the liar, and this makes it more likely that liars experience emotions such as guilt or fear. In addition, lying in these high-stakes contexts is probably more mentally taxing, and liars are more likely to attempt to appear credible in order to avoid getting caught. Therefore, in high-stakes situations nonverbal and verbal cues that reveal emotions, cognitive load, or attempted control may arise in liars. However, the lie detector should be careful not to interpret such cues as indicators of deceit too quickly because truth tellers may also display such cues in high-stakes situations as I discuss in the next guideline.

(3) Consider alternative explanations when interpreting cues of emotions, cognitive load, and attempted control

The difficulty lie detectors face is that not only liars but also truth tellers are more likely to display signs of emotions, cognitive load, and attempted behavioural control in high-stakes situations than in low-stakes situations. When the stakes are high, being disbelieved may also have negative consequences for truth tellers. Just consider how it must feel to be falsely accused by the police of having committed a crime; by an employer about breaking expensive equipment; or by an acquaintance about having an affair. Truth tellers may also experience fear of not being believed in such situations, may have to think hard while responding, and may try to appear credible. Cues of emotions, cognitive load, and attempted control do therefore not conclusively demonstrate that someone is lying, and the lie detector should thus be cautious in interpreting such cues as signs of deceit. Instead, in interpreting emotional responses the lie detector should consider questions such as: "Is my questioning likely to evoke emotions in the respondent, regardless of whether he or she is guilty?", "Is the present situation likely to evoke emotions in the respondent anyway?", "Is this person the type who is likely to be emotional in this situation anyway?" (Ekman, 1985/2001). Similar questions should be considered when interpreting signs of cognitive load or attempted behavioural control (e.g., "Is my questioning likely to evoke cognitive load in the respondent, regardless of whether he or she is guilty?").

(4) Be suspicious but do not show suspicion

People often have the tendency to believe others. They do so sometimes because it is not in their own interest to detect lies (the *ostrich effect*, Chapter 1) or because of a variety of reasons related to the *truth-bias* (Chapter 6). Being suspicious is a necessary prerequisite to catch liars. Being suspicious may further be beneficial because it typically leads to a more fine-grained level of analysis of nonverbal and verbal behaviour (Atkinson & Allen, 1983; Schul, Burnstein, & Bardi, 1996; Schul, Mayo, Burnstein, & Yahalom, 2007). However, it is important that lie detectors do not show their suspicions. Revealing suspicions may make truth tellers feel uncomfortable and this may result in the *Othello error* (Chapter 14). Suspiciousness may also result in escape routes for liars. For example, it could result in them refusing to talk any longer (e.g., "Why should I speak to you? You don't believe me anyway!") or in "counter-attacks". When a wife accuses her husband of concealing an affair, he could respond with "What kind of a relationship do we have if you can't trust me?" Finally, as I explained in Chapter 3, when liars believe that they are not getting away with their lies, they may try to adapt their nonverbal and verbal behaviour in order to appear more credible. It will hamper lie detection if they succeed in doing this.

(5) Do not make up your mind too quickly about whether a person is lying

Judging whether someone is lying is rarely an easy and straightforward task, and it is therefore important to evaluate all the available information. However, once people have made up their minds about the veracity of a message they have the tendency to interpret additional information in such a way that it supports their decision (Chapter 5). As a result, after making up their minds, lie detectors run the risk of failing to notice further important information or of misinterpreting such information. It is thus important to keep an open mind until all available information has been assessed. In Chapter 5 I explained that the tendency to be open-minded depends on situational and personality factors.

(6) Pay attention to the more diagnostic verbal and nonverbal cues to deceit

Chapter 5 revealed that observers often base their veracity decisions on cues that are not diagnostic cues to deception, such as averted gaze or fidgeting. It sounds therefore plausible that observers may become better at discriminating between truths and lies if they are taught to pay attention to more diagnostic cues to deception. Several "training" studies have addressed this issue. Appendix 15.1 provides the results of these studies and summarises via notes the content of the training given to participants (see also Bull, 2004, and Frank & Feeley, 2003, for reviews of training research).

In all training studies, observers were exposed to short videotaped or audiotaped interviews with a number of people who were telling either truths or lies. Generally, three different procedures were used. Some studies used a "focusing procedure" where observers were asked to pay attention to specific cues and ignore others. Others studies used an "information procedure" where observers received information about the actual relationship between certain behaviours and deception. Yet other studies used an "outcome feedback" procedure where each time observers made a decision they were informed about the accuracy of that decision. In all three types of procedures, the performance of these trained participants was then compared with the performance of untrained participants.

As Appendix 15.1 shows, most studies revealed that trained observers were better at distinguishing between truths and lies than untrained observers, regardless of which training method was used. However, these improvements were typically small. On average the untrained observers detected 53.40% of the truths and lies correctly, and the trained observers 57.66%. In other words, people can to a limited extent be trained to become better lie detectors. The small improvements found in research may not necessarily reflect the true potential of teaching people to detect deceit. The training programmes were typically brief and sometimes lasted no more than 15 minutes. Neither did they address the complex nature of lie detection. For example, in studies utilising the information procedure, observers were taught a set of cues that liars typically display. This approach is limited because not all liars will show these specific sets of cues. Moreover, in all these studies the observers were exposed to low-stakes truths and lies, and low-stakes situations do not provide much opportunity to detect deception. It could thus be that training has larger effects if observers are given more sophisticated training and are exposed to truths and lies told in high-stakes situations.

The training studies revealed two more outcomes that are worth discussing. First, Levine, Feeley, McCornack, Hughes, & Harms' (2005) experiment included bogus training groups that were taught cues that are *not* diagnostic cues to deception. They found that sometimes these bogus training groups performed better than the untrained groups, suggesting that the simple act of training, rather than the content of the training, may improve accuracy. To explain these findings Levine *et al.* suggest that the bogus trained observers assessed the messages more critically than the untrained observers.

Second, some studies revealed that trained observers performed worse than untrained observers. This included Kassin and Fong's (1999) experiment where observers were trained to examine the cues taught by the Inbau group and reported in their manual (Inbau, Reid, Buckley, & Jayne, 2001). I explained in Chapter 5 that many of the cues mentioned in that manual are not diagnostic cues to deception. In quite a few other studies where training made observers worse lie detectors (Köhnken, 1987; Vrij, 1994; Vrij & Graham, 1997), the observers were police officers rather than the type of participants typically used in these training studies: undergraduate students. This suggests that students benefit more from these training sessions than do police officers. This became apparent in one of my own studies (Vrij & Graham, 1997), where students performed better as a result of the information they received, whereas police officers performed worse after having received the same information.

I can only speculate as to why police officers do not appear to benefit from the information they are given. One possible explanation is that the information confuses them (see also Köhnken, 1987). Perhaps the information we (Vrij & Graham, 1997) gave about the relationship between personality traits and deceptive behaviour was confusing for police officers who I suspect are not familiar with personality theories. The student observers in our experiment were psychology students and hence familiar with personality theories (albeit not with the relationship between personality traits and deception). A second possible explanation is that police officers refused to use the information provided because they did not believe this information. For example, in my study (Vrij, 1994) the observers were told that liars typically show a decrease in hand and finger movements. This contradicts police officers' beliefs, as they typically assume that an increase in hand and finger movements indicates deception (Chapter 5). Perhaps the officers refused to accept the information provided by an outsider (the experimenter) and continued to rely on their own experience and beliefs instead.

Another approach to enable lie detectors to focus on diagnostic cues to deception is identifying good lie detectors and unravelling the strategies they use. Perhaps their successful strategies can be taught to other lie detectors. At present some researchers are interviewing a selected group of individuals who they believe are exceptionally good lie detectors (O'Sullivan, 2005, 2007; O'Sullivan & Ekman, 2004). The outcomes of this research project have yet not been reported.[3]

(7) Pay attention to nonverbal and verbal cues simultaneously

Observers should pay attention to nonverbal behaviour and speech simultaneously. They can do this in three different ways, and I believe that all three methods will enhance lie detection. First, observers could take into account a mixture of nonverbal and verbal cues, without looking at the relationship between the two sets of cues. For example, they can focus on the types of detail a person mentions as well as on the movements he or she makes, without examining what was said when which movements were made. Our research and that of others has consistently shown that more accurate truth/lie decisions can be made when both nonverbal behaviour and speech content are taken into account than when focusing on nonverbal communication or speech content separately.[4]

Second, observers could examine nonverbal behaviour in relation to speech content, an approach common in communication research,[5] but often ignored by deception researchers. In Chapter 3 I explained the potential of this approach. In one study we were not able to distinguish between truth tellers and liars when we counted the number of illustrators they made, but we did find differences when we related the illustrators to the speech content. We found that truth tellers and liars differed in specific types of illustrators they made during specific parts of their statements (Caso, Maricchiolo, Bonaiuto, Vrij, & Mann, 2006).

Third, observers could examine mismatches between nonverbal behaviour and speech content. As discussed in Chapter 3, strong emotions may result in specific facial expressions. Since certain emotions are

[3]Not everyone is convinced that the exceptional talent of these lie detectors has truly been established. See Bond and Uysal (2007) for a critical article about how these exceptionally good lie detectors were selected.

[4]Ekman & O'Sullivan, 1991; Porter & Yuille, 1995, 1996; Porter, Yuille, & Birt, 2001; Porter, Yuille, & Lehman, 1999; Vrij, Edward, Roberts, & Bull, 2000; Vrij, Evans, Akehurst, & Mann, 2004; Vrij, Akehurst, Soukara, & Bull, 2004a; Vrij & Mann, 2004.

[5]Bavelas & Chovil, 2006; Bavelas, Chovil, Coates, & Roe, 1995; Bavelas & Gerwing, 2007; Freedman, 1972; Kendon, 1994; 2004; McNeill, 1985, 1992.

specifically related to certain patterns of facial expressions, it is likely that the person actually experiences the emotion when a specific facial expression pattern emerges, even when he or she denies feeling the emotion. In this case, the mismatch between speech content and emotional expression gives away that the person is lying about the emotion he or she experiences (see also Ekman, 1985/2001 and Ekman & O'Sullivan, 2006). Other cues, however, are less diagnostic. As I discussed in Chapter 8, some people will display clear signs of distress when they talk about a negative event they have experienced, whereas others do not. In other words, not showing distress during an interview about an upsetting event is not a valid indicator of deceit (although many observers believe it is, Chapters 5 and 8). Therefore, when lie detectors believe that there is a mismatch between someone's nonverbal behaviour and speech content they should be careful about how to interpret this. A final judgment that the person is lying should not be made too easily, and alternative explanations (in the example above, people do not necessarily show signs of distress when discussing upsetting events) should be considered.

(8) Pay attention to deviations from a person's honest reactions in similar situations: the comparable truth

Lie detectors should take interpersonal and intrapersonal differences into account when making veracity judgements. Therefore, the relevant question for the observer to ask is whether the nonverbal behaviour and speech patterns displayed by a person differs from this person's known truthful responses. For this technique to work, it is essential that the known truthful response (e.g., baseline response) is made under similar conditions to the response under investigation (*the comparable truth*, Chapter 3), because only then can intrapersonal differences be taken into account. People respond differently in high-stakes situations than in low-stakes situations (Chapter 3). Therefore, if the response under investigation is made in a high-stakes situation, so should be the baseline response. In addition, personal involvement with the topic also impacts on someone's behaviour (Chapter 3). Thus, if the response under investigation is about a topic the person really cares about, so should the baseline response be. Furthermore, since people show different behaviours when they are interviewed by different people (Chapter 3), it is important that the response under investigation and baseline response are both elicited by the same interviewer.

In Chapter 3 I gave an example of how the comparable truth technique could be employed. During a videotaped real-life police interview, a man suspected and later convicted of murder was asked to describe his activities during a particular day (Vrij & Mann, 2001a). The murder suspect described his activities during the morning (went to work), afternoon (visited a market), and evening (visited a neighbour). Detailed analyses of the videotape revealed a sudden change in behaviour as soon as he started to describe his activities during the afternoon and evening. One possible reason for this may have been that he was lying. Evidence supported this view. Police investigations could confirm his story with regard to his morning activities, but revealed that his statement about the afternoon and evening was fabricated. In reality, he met the victim in the afternoon and killed her later on that day. In this case, we were able to make a good comparison. The man described a seemingly normal day, and there are no good reasons why different behaviours would emerge while describing different parts of that day.

In a second real-life example of employing this technique, I was invited by the police to watch a videotaped interview with a man suspected of kidnapping his wife. The man was asked to describe his activities during a particular day. His descriptions of the four different phone calls he had made drew my attention. He described the first phone call in much detail and included quotes of both what he had said and what the other person had said (e.g., "And I said . . .., and then he said . . ..", etc.). This pattern also emerged when he described his second and fourth phone call of the day. However, the description of the third phone call was different. The man gave less detail and his answer did not include any quotes. When the police asked him to elaborate on this third phone call, he provided some details but quotes still did not emerge. When I pointed out to the police the difference between the third and the other phone calls they told me that they had reasons to believe that the third phone call never took place.

The comparable truth technique has inevitable shortcomings and mistakes will still be made when utilising this technique. The main problem is that it is difficult to rule out that the observed nonverbal and verbal differences are caused by factors other than deceit. Open-mindedness when interpreting the differences in behaviour and speech is thus crucial. Also, differences between the baseline message and the message under investigation may be subtle and therefore difficult to spot.[6] Finally, an absence of behavioural and speech-related differences between the baseline message and the message under investigation does not necessarily mean that the person is telling the truth.

[6] In this respect, recent computer technology developments where the frequency of occurrence of a number of nonverbal behaviours can be counted and displayed in real time may be helpful to observers (Dente, Barath, Ng, Vrij, Mann, & Bull, 2006; Rothwell, Bandar, O'Shea, & McLean, 2006).

(9) Employ indirect lie detection techniques

Another way to improve lie detection skills is by encouraging people to better use their existing potential to distinguish between truths and lies. This can be achieved by asking observers indirectly whether they think someone is lying (e.g., *indirect lie detection*, Buller & Burgoon, 1996; DePaulo, 1994; DePaulo, Charlton, Cooper, Lindsay, & Muhlenbruck, 1997; DePaulo & Morris, 2004; Granhag, 2006; Vrij, 2001). For example, rather than attempting to answer the question "Is the person lying?" observers could attempt to answer the question "Is the person having to think hard?" Research has shown that observers are better lie detectors when answering the latter indirect lie detection question than when answering the former direct lie detection question. In three studies observers could only discriminate between truth tellers and liars by answering the indirect lie detection question, and in all studies liars were perceived as having to think harder than truth tellers (Landström, Granhag, & Hartwig, 2005, 2007; Vrij, Edward, & Bull, 2001b). The findings of our study further revealed that only participants (police officers in this study) who answered the indirect lie detection question paid attention to the cues that actually discriminated between truth tellers and liars on the videotape, such as a decrease in hand and finger movements. In other words, the instruction to look for cues of hard thinking (e.g., cognitive load) directed the observers' attention to more diagnostic cues to deception.

Interviewing to Detect Deception

On many occasions observers have the opportunity to interview the alleged liar. I provide guidelines on how observers could use this opportunity to discriminate between truths and lies. The concept of "interviewing to detect deception" has been ignored by deception researchers who examine nonverbal and speech-related cues to deceit for a long time (although research in this area is now starting to emerge). Furthermore there is little reported about interviewing to detect deceit in police manuals. Those manuals typically discuss cues to deception but do not explain what interviewers could do to increase the likelihood that liars will display such cues. This could explain why in our experiment no clear or effective strategies emerged when we observed how experienced police officers interview to detect deceit (Hartwig, Granhag, Strömwall, & Vrij, 2004).

One interviewing to detect deception protocol that does exist is the Behaviour Analysis Interview (BAI) described by Inbau *et al*. (2001). I discussed the BAI protocol in Chapter 7, and explained my reservations about this technique. In part, these were related to the assumption that when answering certain questions in criminal investigative interviews (e.g., "Did you take the money?") liars will feel less comfortable than truth tellers. I have made clear throughout this book that this cannot be assumed, and that innocent suspects may feel equally concerned and uncomfortable as liars when answering such questions in high-stakes situations.

I do not think that in criminal investigations, or other high-stakes settings, questions can be asked that will necessarily raise more concern in liars than in truth tellers. None of the guidelines that follow therefore aim to raise concern in interviewees. In contrast, I think that it is possible to ask questions that are more difficult to answer for liars than for truth tellers in some circumstances, and some of the guidelines relate to this. Moreover, when observers have some evidence (e.g., fingerprints) they could strategically use that evidence to pinpoint liars, as I will also explain. Because this interviewing to detect deception approach is new, not all recommendations that I give have yet been tested. Where evidence is available I will indicate this in the text by either including the references or referring to previous chapters.[7]

(10) Use an information-gathering interview style

In Chapter 3 I explained that the police commonly use a mixture of two types of interview styles: information gathering and accusatory. In the information-gathering style, interviewers request suspects to give detailed statements about their activities through open questions (e.g., "What did you do yesterday between 3pm and 4pm?"; "You just mentioned that you went to the gym, who else was there?"). By comparison, in the accusatory style, interviewers confront suspects with accusations (e.g., "Your reactions make me think that you are hiding something from me"). Information-gathering interviews encourage suspects to talk, whereas accusatory interviews often yield short denials (e.g., "I am not hiding anything"). Therefore, information-gathering interviews typically elicit more information about an event and result in longer responses than accusatory interviews (Fisher, Brennan, & McCauley, 2002; Vrij, Mann, & Fisher, 2006b; Vrij, Mann, Kristen, & Fisher, 2007).

For several reasons I believe that an information-gathering interview style is desirable for lie detection purposes. First, as I explained in

[7] Many of the guidelines below require interviewees to talk. I believe that interviewees are generally willing to talk even in situations where it may be less expected, such as in police interviews. In their analysis of 1067 audiotaped police interviews, Moston, Stephenson and Williamson (1993) found that only 5% of suspects remained silent.

Chapter 4, a good lie detection strategy is to check the factual information provided by an alleged liar with the available evidence. The more factual information the interviewee provides, which is most likely to happen in information-gathering interviews, the more opportunities for the lie detector to check. Second, information-gathering interviews result in more nonverbal cues to deceit than accusatory interviews (Chapter 3). Longer statements, which are most likely to occur in information-gathering interviews, typically reveal more nonverbal cues to deception than shorter statements, because the more someone talks the more opportunities for nonverbal cues to deception to occur. Additionally, being accused of wrongdoing (i.e., accusatory interview style) is likely to affect the behaviour of both truth tellers and liars in a similar way, and the accusation has a stronger impact on someone's nonverbal behaviour than the act of lying. Consequently, differences in nonverbal behaviour between truth tellers and liars are overshadowed by the accusation (Chapter 3).

The third advantage of employing an information-gathering interview is that it also results in more verbal cues to deceit (Vrij, Mann *et al.*, 2007). The more words included in a statement the more opportunities for verbal cues to deceit to occur, because words are the carriers of such cues. Fourth, information-gathering interviewing does not involve accusing suspects of any wrongdoing or other tactics designed to make suspects feel uncomfortable. It may therefore be a safeguard against false confessions, because innocent suspects do sometimes falsely confess when they feel uncomfortable (Gudjonsson, 2003). Fifth, veracity judgements in accusatory interviews are made with more confidence than those in information-gathering interviews (Vrij, Mann *et al.*, 2007). Since veracity judgements are often inaccurate (Chapter 6), feeling confident about these judgements is problematic. If lie detectors believe that they are not confident enough to make a veracity judgement they defer making such judgements and instead decide to further investigate the case (see also Levine & McCornack, 1992). Perhaps evidence about the veracity of the statement will arise from such further investigations. Finally, several of the interview guidelines discussed below are easier to employ in information-gathering interviews than in accusatory interviews.[8]

[8] Information-gathering interviews are also desirable for an important reason that is not related to deception. Research with murderers and sex offenders has shown that they may result in unprompted confessions (Holmberg & Christianson, 2002). When suspects feel that they are respected and acknowledged, they may gain more confidence and mental space which allows them to admit to criminal behaviour.

(11) Detecting deception may be the easiest during the first interview

Lie detection is perhaps easier during the first interview than during subsequent interviews. During the first interview, liars may not yet be fully prepared. They may not have rehearsed the story that they are going to tell and may not have worked out their answering strategy. Members of the resistance during the Second World War who were caught by the Germans and were subsequently interviewed by German police were often poorly prepared for those interviews (J. Vrij, 2001). All their precaution activities were focused on avoiding capture by the German police, and the equally important aspect of what to say if they were caught was ignored.

In addition, after the initial phase of the interview, interview dynamics could make interpreting the responses of the alleged liar more complex. For example, the interviewer may become suspicious and this suspicion, in turn, may affect the interviewee's behaviour (Chapter 3). Moreover, when people communicate with each other, behavioural matching and behavioural synchrony takes place. People may mirror each other's posture or they may converge in how quickly and loudly they speak. They may also reciprocate each other's gazing, nodding, and smiling behaviour (Chapter 3). Consequently, as the interview progresses, the behaviour of the interviewee may become influenced by the behaviour of the interviewer. This complicates interpreting the behaviour of the interviewee.

(12) Be informed about the factual evidence

Before starting the interview, a lie detector should make himself or herself knowledgeable about the factual evidence of the case. Lie detection could become easier if there is factual evidence and if the lie detector is aware of this evidence. There is always a chance that a liar will contradict the evidence, and these lies can readily be detected if the lie detector is aware of the evidence (Chapter 4). Whether evidence is available also determines what interview technique to employ. In cases where there is no evidence, the lie detector could use the strategies outlined in Guidelines 13 to 16. When there is evidence, the lie detector could also use this evidence strategically, as outlined in Guideline 17.

(13) Let the person repeat themselves

Sometimes liars could be caught if they are asked to repeat themselves. Many liars do not fully prepare themselves (Ekman, 1985/2001), and are taken by surprise by certain questions. To avoid appearing suspicious they then immediately have to fabricate an answer to these unexpected questions. Since liars will not be able to properly rehearse these impromptu fabrications after expressing them, they run the risk of forgetting exactly what they have actually said. Lie detectors could exploit this by asking the person to repeat this information at a later stage in the interview.

Liars who have prepared themselves are likely to remember what they have said before when asked to repeat the information they gave prior in the interview. I think that they could be caught out if the request to repeat the information is asked in a different format than the initial question. For example, if the initial question was "How old are you?" the follow-up question could be "What is your date of birth?" (Walczyk, Schwartz, Clifton, Adams, Wei, & Zha, 2005). I expect liars to contradict themselves more often than truth tellers. I also think that many liars will find this second question difficult to answer and so longer latency periods may occur.

Many more questions could be asked to make interviewees recall their stories from a somewhat different perspective. I will give one more example. Suppose that someone reports that between two and five o'clock in the afternoon he went by car from his home to a shopping centre, visited three shops, had a drink, and drove home again. Temporal questions could subsequently be asked about each of the listed activities ("How long did it take you to get from your home to the first shop?", "For how long were you in the first shop?" etc.). The summation of the total duration of each of these activities should then equal the total time span described in the initial statement (in this example three hours). I expect liars' descriptions of the duration of the individual activities to deviate more from the initially reported total time span than truth tellers' descriptions. Alternatively, interviewees could be asked what they did at a specific time (e.g., "What did you do at four o'clock?"). When comparing this answer to the initial statement, I expect liars' answers to be more inconsistent than truth tellers' answers.

(14) Let interviewees elaborate

Lie detectors could invite interviewees to elaborate on what they previously said in response to the initial open question during the information-gathering interview. This may be difficult, particularly for liars. Liars may have prepared themselves about what they are going to say, but are unlikely to have prepared in as much detail as is required in extensive elaboration requests. They then face a dilemma. They could make things up spontaneously, but this is sometimes difficult. Alternatively, they could say that they do not remember anything else. This strategy may look suspicious if they have given a detailed account in their original answer. People rarely spontaneously recall all the details they know. Therefore, initially providing lots of details followed by a total silence will look suspicious, and I therefore expect many liars to provide additional detail. This will benefit the lie detector, as I explained in Guideline 10. The elaborative answer could contain additional factual information that the lie detector can check (see also Schweitzer & Croson, 1999). The elaborative answer may also provide nonverbal and verbal cues to deception.

An alternative way to obtain elaborations is by asking the interviewee *surprise questions*. These are questions that the interviewee has not anticipated, and that he or she cannot answer with "I don't know" without raising suspicion. Interviewees have therefore no option other than to provide spontaneous (unprepared) answers to these questions. This may be difficult for liars, and the answers may thus include nonverbal and verbal signs to deceit. Which questions are surprise questions depends on the situation and the interviewee, but if, for example, the interviewee says that she went to a restaurant last night, a surprise question could be "What "specials" did they have on last night"? Or, if someone describes his job, surprise questions could be "What qualifications are required to get into your job"?, "How many days annual leave do you get?" or "Do you have to take them at a specific time of year?"

(15) Ask temporal questions when a scripted answer is suspected

A good strategy for liars is to fabricate a story which is in fact true, but which happened at a time other than when the liar claims (*embedded lie*, Chapters 8, 9 and 14). For example, a guilty suspect who denies involvement in the crime could claim that he was at the gym at the particular time the crime took place. If he is familiar with that gym then he can now truthfully recall an experience at the gym, can describe its layout, the equipment that he uses there, and so on. The only fabricated part in this story is *when* he was there. Lie detectors should be aware of this lying strategy. Questions about the layout of the gym and activities whilst there are not effective. Instead, questions should be asked that are specifically related to the particular time that the interviewee claims to have been there, because the guilty suspect is lying about this aspect. For example, the interviewer could ask which instructor was present at the time he or she claims to have visited the gym, and who else was there.

Let me explain how this strategy could have been used in interviewing Maxine Carr. She was the girlfriend of Ian Huntley, the man who was convicted of murdering two 10-year-old girls in Southern

England on Sunday 4 August 2002 (I discussed the Ian Huntley case in Chapter 3). Maxine Carr told an embedded lie when covering up for her boyfriend. Although she had been miles from home, she told the police that she was with Huntley at home all day on the Sunday that the girls went missing. She hereby gave Huntley an alibi. When the police asked Carr what she had done on that Sunday, she answered in some detail that they had prepared and consumed a roast lunch. I expect that Carr and Huntley have spent a Sunday in this way before and that she thus described a truly experienced sequence of events, although these activities did not take place on the day she claimed they took place. In the parts of the interview with Carr that I have seen on television no further questions were asked about their activities on that Sunday. The police, however, could have asked Carr questions about her experiences when shopping for the ingredients for this lunch (e.g., when and where she did that shopping, the weather at the time of the shopping, how many people were in the shops, who served her in these shops, how she paid). I expect that she would have found those questions difficult to answer. She probably would have recalled a memory of another time that she went shopping for a Sunday lunch, pretending that it was the shopping for the Sunday 4 August lunch. However, if she did this, she would also have realised that the circumstances on the day she recalled may have differed from the circumstances on the day she claimed to have been shopping for the lunch on 4 August.

(16) Make the interviews more cognitively demanding

In situations when lying is cognitively more demanding than truth telling, lie detectors could exploit the increased mental demand experienced by liars to elicit more cues to deceit. To answer the question of *when* lying is more cognitively demanding than truth telling, one should take into account *why* lying would be more cognitively demanding than truth telling. In Chapter 3 I explained that six aspects of lying contribute to an increase in mental load. First, formulating the lie itself may be cognitively demanding. A liar needs to invent a story and must monitor their fabrication so that it is plausible and adheres to everything the observer knows or might find out. In addition, liars must remember their earlier statements, and know what they told to whom in order to maintain consistency. Liars should also avoid making slips of the tongue, and should refrain from providing new leads.

A second aspect of lying that adds to mental load is that liars are typically less likely than truth tellers to take their credibility for granted. As such, liars will be more inclined than truth tellers to monitor and control their demeanour so that they will appear honest to the lie detector. Monitoring and controlling their own demeanour should add cognitive demand for liars. Third, because liars do not take credibility for granted, they may monitor the *interviewer's* reactions more carefully in order to assess whether they are getting away with their lie. Carefully monitoring the interviewer also requires cognitive resources.

Fourth, liars may be preoccupied by the task of reminding themselves to act and role-play, which requires extra cognitive effort. Fifth, liars have to suppress the truth while they are lying and this is also cognitively demanding. Finally, while activation of the truth often happens automatically, activation of the lie is more intentional and deliberate, thus requiring mental effort.

These six reasons as to *why* lying is more cognitively demanding could give us insight into *when* it is more cognitively demanding. That is, lying is more cognitively demanding to the degree that these six principles are in effect. I believe that for at least some of these six principles to be fulfilled, two elements are required. First, lying is likely to be more demanding than truth telling only *when interviewees are motivated to be believed*. Only under those circumstances can it be assumed that liars take their credibility less for granted than truth tellers and hence will be more inclined than truth tellers to monitor their own nonverbal and verbal behaviour and/or the interviewer's reactions. Second, for lying to be more cognitively demanding than truth telling, *liars must be able to retrieve their truthful activity easily and have a clear image of it*. Only when liars' knowledge of the truth is easily and clearly accessed will it be difficult for them to suppress the truth. On the other side of the equation, *truth tellers also need to have easy access to the truth* for their task of truthfully reporting an event to be relatively undemanding. If truth tellers have to think hard to remember the event in question (e.g., because it was not distinctive or it occurred long ago), their cognitive demands may exceed the cognitive demands that liars require for fabricating a story.

In high-stakes situations interviewees will typically be motivated to be believed, and therefore this criterion will most often be fulfilled. Whether truth tellers and liars have easy and clear access to the truth depends on several factors, including the distinctiveness of the event in question and the time span between the event and the interview. Interviewers could ask interviewees whether they have a clear memory of the relevant activities before asking questions about them. Only when interviewees indicate that they do have good recollection, should the guideline outlined in the next paragraphs be employed.

A lie detector could exploit the differential levels of cognitive load that truth tellers and liars experience to discriminate more effectively between them. Liars who require more cognitive resources than truth tellers for the act of storytelling will have fewer cognitive resources left over than truth tellers. This makes liars vulnerable, and so if cognitive demand is further raised, which could be achieved by making additional requests (some suggestions for which follow), liars may not be as good as truth tellers in coping with these additional requests. Liars' lack of coping could impair their story telling, could result in ineffectively complying with the additional request, or in both. Due to liars' keenness to be believed, it sounds plausible that they will employ their cognitive resources to the storytelling aspect and therefore predominantly differences between truth tellers and liars in dealing with the additional request will occur (e.g., liars will comply less with the additional request than truth tellers). However, as the importance of complying with the additional request increases, this strategy of "neglecting" the additional request will become less viable, and, consequently, liars' story telling will be impaired.

There are numerous additional requests that lie detectors could make, and I will introduce three of them. Lie detectors could ask interviewees to recall their activities in reverse order. We demonstrated the benefit of this instruction for the purpose of lie detection in an experimental study (Vrij, Mann, Fisher, Leal, Milne, & Bull, in press). Truth tellers took part in a staged event and recalled this event in a subsequent interview. In contrast, liars did not take part in the event but pretended that they had in the subsequent interview. Truth tellers and liars were either instructed to recall the event in reverse order or were just asked to give their story as they saw fit (e.g., control condition). The reverse order condition resulted in interviewees describing first what happened at the end of the event, then what happened just before that, and so on. The control condition resulted in interviewees describing their activities in normal chronological order.

We found that truth tellers were more successful than liars in complying with the request to tell their stories in reverse order. We further found that many more nonverbal and verbal cues differentiated truth tellers from liars in storytelling in the reverse order condition (eight) than in the control condition (one) and most of the cues that distinguished truth tellers from liars in the reverse order condition were signs of cognitive load. For example, in the reverse order condition, liars produced more speech hesitations and speech errors and spoke slower than truth tellers, whereas in the chronological order interviews no differences in hesitations, errors and speech rate emerged between truth tellers and liars. In other words, the introduction of an additional request resulted in liars performing worse than truth tellers in both complying with the additional request and in storytelling.

We then showed police officers video clips of the interviews with either the participants who recalled their story in chronological order or with the participants who recalled their story in reverse order. The police officers who saw the chronological order interviews classified only 50% of the truths and 42% of the lies correctly (total accuracy rate was 46%). The police officers who saw the reverse order interviews performed remarkably better and classified 56% of the truths and 60% of the lies correctly (58% total accuracy). In itself, a 60% lie detection accuracy in the reverse order interviews may not appear high, but an increase from 42% (chronological order interviews) to 60% (reverse order interviews) is substantial. There is also anecdotal evidence to support the reverse order technique. Investigators who use this technique indicated that suspects sometimes give themselves away by their obviously non-credible stories that are often riddled with inconsistencies.

A second additional request that could be introduced is asking interviewees to carry out a secondary task at the same time as recalling their stories. For example, interviewees could be asked to recall their stories while conducting a computer driving simulation task.[9] This means that interviewees would have to divide their attention between the storytelling and driving task. Due to the additional resources that are needed for telling the lie, liars may find this dual task more cognitively difficult than truth tellers and may not cope as well. Due to liars' keenness to be believed, it sounds plausible that they will focus their attention more on the storytelling aspect and therefore predominantly differences between truth tellers and liars in the driving task will occur (e.g., liars will perform worse at the driving task than truth tellers). However, if the salience of complying with the additional request increases (i.e., interviewees are informed that mistakes on the task are viewed as "suspicious"), this strategy of "neglecting" the additional request will become less viable, and, consequently, liars' storytelling may become impaired.

An experiment with children reveals a third type of additional request that can be made: asking event-irrelevant questions (Quas, Davis, Goodman, & Myers, 2007). Children played individually with a male confederate who touched each child twice on their stomach, nose, and neck. In the subsequent interview children were asked to tell the truth or lie when asked questions about body touch. They were also asked a series of questions about the event that were unrelated to body touch and were asked to answer those questions truthfully. The children who

[9] Safety considerations prevent me from recommending using this method with real-life driving.

lied about the body touch answered these unrelated questions less accurately than the children who told the truth about the body touch. Apparently, remembering and rehearsing the lie required cognitive resources, and by devoting their resources to the lie children had difficulty in conducting an adequate memory search for other event details.

(17) If evidence is available, use it strategically

Suppose someone left his briefcase in a bookshop on top of a box of stationery and that, when he returned to the shop to collect his briefcase, he noticed that his wallet had been stolen from the briefcase. Further suppose that the police found fingerprints on the briefcase that did not belong to the wallet's owner but did belong to another customer who had visited the bookshop. This makes the customer a suspect but not necessarily the culprit of the theft, because an innocent explanation for the fingerprints is also possible: perhaps the customer moved the briefcase to look in the box of stationery. In such circumstances the police need to interview the suspect to find out the truth. Traditionally they employ interview styles where they present the evidence at the beginning of the interview ("Your fingerprints have been found on the briefcase") (Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Leo, 1996a). This interview technique is limited, the problem being that it gives guilty suspects the opportunity to fabricate a story that is consistent with the evidence. The Strategic Use of Evidence (SUE) technique, developed by Swedish researchers, is an interview method that makes it less likely that liars will be consistent with the available evidence.[10]

The first step of the SUE technique is to ask the suspect to describe his or her activities, hence in the example above, to describe his or her activities in the bookshop. It is hereby important not to reveal the fingerprint evidence. Research concerning the strategies truth tellers and liars use when they are interviewed provide insight into what kind of recall they are likely to give (Colwell, Hiscock-Anisman, Memon, Woods, & Michlik, 2006; Granhag & Strömwall, 2002; Granhag, Strömwall, & Hartwig, 2007; Hartwig, Granhag, & Strömwall, 2007; Strömwall, Granhag, & Landström, 2007). This research shows that truth tellers tend to "keep the story real" and "to tell the truth like it happened". In contrast, liars' strategy is to avoid mentioning possibly incriminating

[10] Granhag & Hartwig, 2007; Granhag & Strömwall, 2002; Granhag, Strömwall & Hartwig, 2007; Hartwig, Granhag, & Strömwall, 2007; Hartwig, Granhag, Strömwall, & Kronkvist, 2006; Hartwig, Granhag, & Vrij, 2005; Strömwall, Hartwig, & Granhag, 2006. The SUE technique resembles Van den Adel's (1997) "blocking escape routes" technique, an interview method taught in the Netherlands and outlined in English in Vrij (2003).

evidence. In the example above, it is thus more likely that truth tellers will mention the briefcase than liars. Truth tellers have nothing to hide and will recall what has happened and this included touching the briefcase; liars do not wish to associate themselves with the crime they have committed and thus distance themselves from the briefcase.

However, not mentioning touching the briefcase still does not indicate guilt. Therefore, further questioning is required after the free recall. In the second phase of the SUE technique, the questioning phase, the interviewer asks questions, including about the briefcase without revealing the incriminating fingerprint evidence. Another liars' strategy is to deny having incriminating knowledge when asked about it (Granhag & Strömwall, 2002; Granhag *et al.*, 2007; Hartwig *et al.*, 2007). Hence, there is a chance that a liar will deny having touched the briefcase, and thereby contradicts the evidence known to the lie detector. The third phase of the SUE technique is to reveal the evidence and in case contradictions between the evidence and statement did emerge, to ask the suspect to explain these contradictions.

Hartwig *et al.* (2006) tested the SUE technique in their experiment, using the wallet taken out of the briefcase scenario described above. Half of the participants were instructed to steal the wallet from the briefcase, the other half were instructed to buy something that was in the box of stationery underneath the briefcase. Swedish police trainees interviewed the mock suspects. Prior to the interviews, half of the interviewers were trained how to use the SUE technique and were asked to use this technique in the subsequent interview. The other half of the interviewers did not receive training and were instructed to interview the suspects in the manner of their own choice. The untrained interviewers obtained 56.1% accuracy rate which is similar to that typically found in nonverbal and verbal deception detection research (Chapter 6). SUE trained interviewers, however, obtained 85.4% accuracy rate and this is the highest accuracy rates ever reported in this type of deception detection research.[11] Further analyses showed that guilty suspects contradicted the evidence more than innocent suspects, but importantly, particularly when they were interviewed by SUE-trained interviewers.

[11]The fact that the participants were police trainees rather than experienced police detectives cannot explain why the accuracy rate in the control condition was low. In a study that resembled the SUE experiment (Hartwig, Granhag, Strömwall, & Andersson, 2004), experienced police detectives participated, and, similar to the SUE experiment, these officers (untrained in SUE) could freely interview the suspects. These experienced officers did not fare better (57% accuracy rate) than the trainees in the control condition of the SUE experiment (56% accuracy rate).

Appendix 15.1 The effect of training observers on accuracy scores: only total accuracy scores (truth and lie accuracy combined) are given

	Control (no training)	Training 1	Training 2	Training 3	Training 4
DeTurck (1991, 183 students)	54%	69%[1]			
DeTurck, Feeley, & Roman (1997, 165 students)	57%	55%[2]	62%[3]	64%[4]	
DeTurck, Harszlak, Bodhorn, & Texter (1990, 188 students)	55%	64%[5]			
DeTurck & Miller (1990, 390 students)	53%	61%[6]			
Fiedler & Walka (1993, 72 students)	53%	65%[7]	65%[8]		
Kassin & Fong (1999, 40 students)	56%	46%[9]			
Köhnken (1987, 80 police detectives)	47%	42%[10]	48%[11]		
Levine, Feeley *et al.* (2005, exp. 1, 256 students)	52%	49%[12]	56%[13]		
Levine, Feeley *et al.* (2005, exp. 2, 90 students)	53%	53%[14]	52%[15]		
Levine, Feeley *et al.* (2005, exp. 4, 158 students)	50%	58%[16]	56%[17]		
Porter, Woodworth, & Birt (2000, 20 parole officers and 95 students)	52%	61%[18]	60%[19]	59%[20]	
Porter, Woodworth, McCabe *et al.* (2007, 151 students)	53%	49%[21]	49%[22]		
Santarcangelo, Cribie, & Ebesu Hubbard (2004, 97 students)	65%	68%	66%[23]	66%[24]	72%[25]
Vrij (1994, 360 police detectives)					
One interview	49%	52%[26]	54%[27]		
Two interviews spontaneously, total image[28]	51%	47%[29]	47%[30]		
Two interviews spontaneously, hands only[31]	44%	56%[32]	60%[33]		
Vrij & Graham (1997, 40 students)	42%	55%[34]			
Vrij & Graham (1997, 29 police officers)	54%	48%[35]			
Zuckerman, Koestner, & Alton (1984, 132 students)	62%	70%[36]			
Zuckerman, Koestner, & Colella (1985, 117 students)					
Face only[37]	53%	59%			
Speech only[39]	61%	63%[40]			
Face plus speech[41]	59%	65%[42]			

Notes: [1]Observers were asked to pay attention to message duration, response latency, pauses, nonfluencies, self adaptors and hand gestures; [2]Observers were asked to pay attention to speech errors, pauses, response latency and message duration; [3]Observers were asked to pay attention to self adaptors, hand gestures, head movements and hand movements; [4]Observers were asked to pay attention to speech errors, pauses, self adaptors, and hand gestures; [5]Observers were asked to pay attention to message duration, response latency, pauses, nonfluencies, self adaptors and hand gestures; [6]Observers were asked to pay attention to message duration, response latency, pauses, nonfluencies, self adaptors and hand gestures; [7]Observers were given information about the relationship between deception and smiles, head movements, self adaptors, pitch of voice, speech rate, pauses and channel discrepancies; [8]Observers were given information described in note 7 plus outcome feedback; [9]Observers were shown two videotape segments from the *Reid Technique: Interviewing and interrogation;* [10]Observers were asked to pay attention to changes in head movements, eyeblinks, gaze, illustrators, self adaptors, body movements and foot and leg movements; [11]Observers were asked to pay attention to changes in speech rate, pauses, speech errors and speech hesitations; [12]Observers were told that liars exhibit longer response latencies, more self adaptors, more speech errors and more pauses (valid training group); [13]Observers were told that liars exhibit less eye contact, talk faster, make more postural shifts, and make more foot movements than truth tellers (bogus training group); [14]Observers were told that liars exhibit longer response latencies, more self adaptors, more speech errors and more pauses (valid training group); [15]Observers were told that liars exhibit less eye contact, talk faster, make more postural shifts, and make more foot movements than truth tellers (bogus training group); [16]Observers were told that liars exhibit shorter response latencies, fewer speech errors, less pauses and more foot movements than truth tellers (valid training group); [17]Observers were told that liars show more gaze aversion, more self adaptors, more postural shifts, and talk faster than truth tellers (bogus training group); [18]Observers participated in a training programme including myth dissolution, information about verbal and nonverbal cues to deception, and were given outcome feedback; [19]Observers were given an handout providing a brief overview of the training programme given to the observers in note 18, and were given outcome feedback; [20]Observers were given outcome feedback; [21]Observers were given outcome feedback; [22]Observers were given incorrect outcome feedback; [23]Observers were told that liars display more self adaptors, hand gestures, foot and leg movements and postural shifts than truth tellers; [24]Observers were told that liars produce shorter answers, more pauses, more speech errors and longer response latencies than truth tellers; [25]Observers were told that the answers of liars sound less plausible, less concrete, less consistent and less clear than truth tellers' answers; [26]Observers were given information about the relationship between deception and hand and finger movements; [27]Observers were given the information described in note 26 plus outcome feedback; [28]Observers were exposed to a truthful and deceptive account simultaneously (i.e., they saw the two clips at the same time). The whole participant was visible; [29]Observers were given information about the relationship between deception and hand and finger movements; [30]Observers were given the information described in note 29 plus outcome feedback; [31]Observers were exposed to a truthful and deceptive account simultaneously (i.e., they saw the two clips at the same time). Only the hands of the participants were visible; [32]Observers were given information about the relationship between deception and hand and finger movements; [33]Observers were given the information described in note 32 plus outcome feedback; [34]Observers were given information about the relationship between deception and hand and finger movements for people high or low in public self-consciousness and for good and poor actors; [35]Observers were given information about the relationship between deception and hand and finger movements for people high or low in public self-consciousness and for good and poor actors; [36]Observers were given outcome feedback. Only the "feedback after eight senders" condition is reported; [37]Observers only saw the face of the senders; [38]Observers were given outcome feedback; [39]Observers could only hear the senders; [40]Observers were given outcome feedback; [41]Observers could see and hear the senders; [42]Observers were given outcome feedback.

Epilogue

I started this book by mentioning that the importance of catching liars has increased due to the increase in security threats and terrorist attacks. I also reported that researchers have responded to this demand by testing all sorts of lie detection techniques. I mentioned that we need to examine these techniques carefully. We need to find out how these techniques work and indeed, whether they work, because introducing fallible tools will not bring the solution of catching liars any closer. In this book I reported that several researchers have claimed to have developed techniques that discriminate between truths and lies with very high accuracy. My advice to them is to keep their feet firmly at the ground. In my view no tool is infallible and each tool developed to date has considerable problems and limitations. I hope that this book has convinced readers why I have come to this somewhat pessimistic conclusion. This does not mean that distinguishing between truths and lies is not possible. On the contrary, there are numerous ways and opportunities to pinpoint liars. I hope that this book provides readers with insight into how to achieve this.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner, Spring*, 7–11.

Abe, N., Suzuki, M., Tsukiura, T., Mori, E., Yamaguchi, K., Itoh, M., & Fujii, T. (2006). Dissociable roles of prefrontal and anterior cingulated cortices in deception. *Cerebral Cortex, 16*, 192–199.

Abootalebi, V., Mordai, M. H., & Khalilzadeh, M. A. (2006). A comparison of methods for ERP assessment in a P300-based GKT. *International Journal of Psychophysiology, 62*, 309–320.

Adams, S. H. (1996). Statement analysis: What do suspects' words really reveal? *FBI Law Enforcement Bulletin, October*, 12–20.

Ahlmeyer, S., Heil, P., McKee, B., & English, K. (2000). The impact of pornography on admissions of victims and offenses in adult sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 12*, 123–139.

Akehurst, L., Bull, R., Vrij, A., & Köhnken, G. (2004). The effects of training professional groups and lay persons to use Criteria-Based Content Analysis to detect deception. *Applied Cognitive Psychology, 18*, 877–891.

Akehurst, Köhnken, G., & Höfer, E. (2001). Content credibility of accounts derived from live and video presentations. *Legal and Criminological Psychology, 6*, 65–83.

Akehurst, L. Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology, 10*, 461–471.

Akehurst, L., & Vrij, A. (1999). Creating suspects in police interviews. *Journal of Applied Social Psychology, 29*, 192–210.

Allen, J. J. B., & Iacono, W. G. (1997). A comparison of methods for the analysis of event-related potentials in decetpion detection. *Psychophysiology, 34*, 234–240.

Allen, J. J. B., Iacono, W. B., & Danielson, K. D. (1992). The identification of concealed memories using the event-related potential and implicit behavioural measures: A methodology for prediction in the face of individual differences. *Psychophysiology, 29*, 504–522.

Allen, N. B., & Gilbert, P. (2000). Social intelligence, deception, and psychopathology. In P. Gilbert & K. G. Bailey (2000), *Genes on the couch* (pp. 151–175). Brunner-Routledge.

Allen, V. L. & Atkinson, M. L. (1978). Encoding of nonverbal behavior by high-achieving and low-achieving children. *Journal of Educational Psychology*, *70*, 298–305.

Allwood, C. M., & Granhag, P. A. (1999). Feelings of confidence and the realism of confidence judgments in everyday life. In P. Juslin & H. Montgomery (Eds.), *Judgment and decision making: Neo-Brunswikian and process-tracing approaches* (pp. 123–146). Mahwah NJ: Erlbaum.

Alonso-Quecuty, M. L. (1991). *Post-event information and reality-monitoring: When the witness cannot be honest*. Paper presented at the First Spanish and British Meeting on Psychology, Law and Crime in Pamplona, Spain.

Alonso-Quecuty, M. L. (1992). Deception detection and Reality Monitoring: A new answer to an old question? In F. Lösel, D. Bender, & T. Bliesener (Eds.), *Psychology and Law: International perspectives* (pp. 328–332). Berlin: de Gruyter.

Alonso-Quecuty, M. L. (1996). Detecting fact from fallacy in child and adult witness accounts. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, Law, and Criminal Justice: International developments in research and practice* (pp. 74–80). Berlin: de Gruyter.

Alonso-Quecuty, M. L., Hernandez-Fernaud, E., & Campos, L. (1997). Child witnesses: Lying about something heard. In S. Redondo, V. Garrido, J. Perez, & R. Barbaret (Eds.), *Advances is psychology and law* (pp. 129–135). Berlin: de Gruyter.

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behaviour: Judgmental accuracy from thin slices of the behavioural stream. *Advances in Experimental Social Psychology*, *32*, 201–271.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behaviour as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*, 256–274.

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, *39*, 1037–1049.

Anderson, D. E., Ansfield, M. E., & DePaulo, B. M. (1999). Love's best habit: Deception in the context of relationships. In P. Philippot, R. S. Feldman, & E. J. Coats (Eds.), *The social context of nonverbal behavior* (pp. 372–409). Cambridge, England: Cambridge University Press.

Anderson, D. E., DePaulo, B. M., & Ansfield, M. E. (2002). The development of deception detection skill: A longitudinal study of same-sex friends. *Personality and Social Psychology Bulletin*, *28*, 536–545.

Anderson, D. E., DePaulo, B. M., Ansfield, M. E., Tickle, J. J., & Green, E. (1999). Beliefs about cues to deception: Mindless stereotypes or untapped wisdom? *Journal of Nonverbal Behaviour*, *23*, 67–89.

Anolli, L., Balconi, M., & Ciceri, R. (2003). Linguistic styles in deceptive communication: Dubitative ambiguity and elliptic eluding in packaged lies. *Social Behavior and Personality*, *31*, 687–710.

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

Anson, D. A., Golding, S. L., & Gully, K. J. (1993). Child sexual abuse allegations: Reliability of criteria-based content analysis. *Law and Human Behavior*, *17*, 331–341.

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, *37*, 715–727.

Argo, J. J., White, K., & Dahl, D. W. (2006). Social comparison theory and deception in the interpersonal exchange of consumption information. *Journal of Consumer Research*, *33*, 99–108.

Arnett Jensen, L., Arnett, J. J, Feldman, S. S., & Cauffman, E. (2004). The right to do wrong: Lying to parents among adolescents and emerging adults. *Journal of Youth and Adolescence*, *33*, 101–112.

Arntzen, F. (1970). *Psychologie der Zeugenaussage*. Göttingen, Germany: Hogrefe.

Arntzen, F. (1982). Die Situation der Forensischen Aussagenpsychologie in der Bundesrepublik Deutschland. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 107–120). Deventer, the Netherlands: Kluwer.

Arntzen, F. (1983). *Psychologie der Zeugenaussage: Systematik der Glaubwürdigkeitsmerkmale*. München, Germany: C. H. Beck.

Aron, A., Dutton, D. G., Aron, E. N., & Iverson, A. (1989). Experiences of falling in love. *Journal of Social and Personal Relationships*, *6*, 243–257.

Ask, K. (2006). *Criminal investigation: Motivation, emotion and cognition in the processing of evidence*. PhD-thesis, University of Gothenburg, Sweden, Psychology Department.

Ask, K., & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, *2*, 43–63.

Ask, K., & Granhag, P. A. (2007). Motivational bias in criminal investigators' judgments of witness reliability. *Journal of Applied Social Psychology*, *37*, 561–591.

Atkinson, M. L., & Allen, V. L. (1983). Perceived structure of nonverbal behaviour. *Journal of Personality and Social Psychology*, *45*, 458–463.

Atoum, A. O., & Al-Simadi, F. A. (2000). The effect of presentation modality on judgements of honesty and attractiveness. *Social Behaviour and Personality*, *28*, 269–278.

Aune, R. K., Levine, T. R., Ching, P. U., & Yoshimoto, J. M. (1993). The influence of perceived source reward value on attributions of deception. *Communication Research Reports*, *10*, 15–27.

Aune, R. K., Levine, T. R., Park, H. S., Asada, K. J. K., & Banas, J. A. (2005). Tests of theory of communicative responsibility. *Journal of Language and Social Psychology*, *24*, 358–381.

Backbier, E., Hoogstraten, J., & Meerum Terwogt-Kouwenhoven, K. (1997). Situational determinants of the acceptability of telling lies. *Journal of Applied Social Psychology*, *27*, 1048–1062.

Backbier, E., & Sieswerda, S. (1997). Wanneer en waarom liegen we eigenlijk? *Nederlands Tijdschrift voor de Psychologie*, *52*, 255–264.

Backster, C. (1962). Methods of strengthening our polygraph technique. *Police*, *6*, 61–68.

Backster, C. (1963). The Backster chart reliability rating method. *Law and Order*, *1*, 63–64.

Bagley, J., & Manelis, L. (1979). Effect of awareness on an indicator of cognitive load. *Perceptual and Motor Skills*, *49*, 591–594.

Baldry, A. C., & Winkel, F. W. (1998). Perceptions of the credibility and eventual value of victim and suspect statements in interviews. In J. Boros, I.

Munnich, & M. Szegedi (Eds.), *Psychology and criminal justice: International review of theory and practice* (pp. 74–82). Berlin: de Gruyter.

Baldry, A. C., Winkel, F. W., & Enthoven, D. S. (1997). Paralinguistic and nonverbal triggers of biased credibility assessments of rape victims in Dutch police officers: An experimental study of 'nonevidentiary' bias. In S. Redondo, V. Garrido, J. Perze, & R. Barbaret (Eds.), *Advances in psychology and law* (pp. 163–174). Berlin: de Gruyter.

Ball, C. T., & O'Callaghan, J. (2001). Judging the accuracy of children's recall: A statement-level analysis. *Journal of Experimental Psychology: Applied*, *7*, 331–345.

Barland, G. H. (1984). Standards for the admissibility of polygraph results as evidence. *University of West Los Angeles Law Review*, *16*, 37–54.

Barland, G. H. (1988). The polygraph test in the USA and elsewhere. In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 73–96). London: Sage.

Barnett, M. A., Bartel, J. S., Burns, S., Sanborn, F. W., Christensen, N. E., & White, M. M. (2000). Perceptions of children who lie: Influence of lie motive and benefit. *Journal of Genetic Psychology, 161*, 381–383.

Barrett, E. C. (2005). Psychological research and police investigations: Does the research meet the needs? In L. Alison (Ed.), *The forensic psychologist's casebook* (pp. 47–67). Devon, England: Willan.

Bartholomew, K. & Horowitz, L. M. (1991). Attachment styles among young adults: A test of the four-category model. *Journal of Personality and Social Psychology, 61*, 226–244.

Bashore, T. R., & Rapp, P. E. (1993). Are there alternatives to traditional polygraph procedures? *Psychological Bulletin*, *113*, 3–22.

Baskett, G. D., & Freedle, R. O. (1974). Aspects of language pragmatics and the social perception of lying. *Journal of Psycholinguistic Research*, *3*, 117–131.

Bauer, L. O., Goldstein, R., & Stern, J. A. (1987). Effects of information processing demands on physiological response patterns. *Human Factors,* 29, 213–234.

Bauer, L. O., Strock, B. D., Goldstein, R., Stern, J. A., & Walrath, J. C. (1985). Auditory discrimination and the eye blink. *Psychophysiology*, *22*, 629–635.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*, 3–26.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*, 610–620.

Baumeister, R. F., Hutton, D. G., & Tice, D. M. (1989). Cognitive processes during deliberate self-presentation: How self-presenters alter and misinterpret the behavior of their interaction partners. *Journal of Experimental Social Psychology*, *25*, 59–78.

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *Journal of Personality and Social Psychology*, *16*, 361–383.

Bavelas, J., Black, A., Chovil, N., & Mullett, J. (1990). *Equivocal communication*. Newbury Park, CA: Sage.

Bavelas, J. B. & Chovil, N. (2006). Nonverball and verbal communication: Hand gestures and facial displays as part of language use in face-to-face dialogue. In V. Manusov & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 97–115). Thousand Oaks, CA: Sage.

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, *21*, 394–405.

Bavelas, J., & Gerwing, J. (2007). Conversational hand gestures and facial displays in face-to-face dialogue. In K. Fiedler (Ed.), *Frontiers of social psychology: Social communication* (pp. 283–307). New York: Psychology Press.

Bell, B. E., & Loftus, E. F. (1988). Degree of detail of eyewitness testimony and mock juror judgments. *Journal of Applied Social Psychology*, *18*, 1171–1192.

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

Bell, K. L., & DePaulo, B. M. (1996). Liking and lying. *Basic and Applied Social Psychology*, *18*, 243–266.

Bello, R. (2006). Causes and paralinguistic correlates of interpersonal equivocation. *Jorunal of Pragmatics*, *38*, 1430–1441.

Ben-Shakhar, G. (1991). Clinical judgement and decision making in CQT polygraphy: A comparison with other pseudoscientific applications in psychology. *Integrative Physiological and Behavioral Science*, *26*, 232–240.

Ben-Shakhar, G. (2002). A critical review of the Control Question Test (CQT). In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 103–126). San Diego, CA: Academic Press.

Ben-Shakhar, G., Bar-Hillel, M., & Kremnitzer, M. (2002). Trial by polygraph: Reconsidering the use of the guilty knowledge technique in court. *Law and Human Behavior*, *26*, 527–541.

Ben-Shakhar, G., Bar-Hillel, M., & Lieblich, I. (1986). Trial by polygraph: Scientific and juridicial issues in lie detection. *Behavioural Sciences and the Law*, *4*, 459–479.

Ben-Shakhar, G., & Dolev, K. (1996). Psychophysiological detection through the Guilty Knowledge Technique: Effects of mental countermeasures. *Journal of Applied Psychology*, *81*, 273–281.

Ben-Shakhar, G., & Elaad, E. (2002). The guilty knowledge test (GKT) as an application of psychophysiology: Future prospects and obstracles. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 87–102). San Diego, CA: Academic Press.

Ben-Shakhar, & Elaad, E. (2003). The validity of psychophysiological detection of information with the guilty knowledge test: A meta-analytic review. *Journal of Applied Psychology*, *88*, 131–151.

Ben-Shakhar, G., & Furedy, J. J. (1990). *Theories and applications in the detection of deception*. New York: Springer-Verlag.

Ben-Shakhar, G., Gati, I., & Salamon, N. (1995). Generalization of the orienting response to significant stimuli: The roles of the common and distinctive stimulus components. *Psychophysiology*, *32*, 36–42.

Ben-Shakhar, G., Lieblich, I., & Bar-Hillel, M. (1982). An evaluation of polygraphers' judgments: A review from a decision theoretic perspective. *Journal of Applied Psychology*, *67*, 701–713.

Benz, J. J., Anderson, M. K., & Miller, R. L. (2005). Attributions of deception in dating situations. *The Psychological Record*, *55*, 305–314.

Berliner, L., & Conte, J. R. (1993). Sexual abuse evaluations: Conceptual and empirical obstacles. *Child Abuse and Neglect*, *17*, 111–125.

Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, *14*, 317–335.

Blair, J. P. (1998). *Detecting deception: The effects of Reid Behavioural Analysis Interview training*. Masters thesis, Western Illinois University.

Blair, J. P., & Kooi, B. (2004). The gap between training and research in the detection of deception. *International Journal of Police Science and Management*, *6*, 77–83.

Blair, T. M., Nelson, E. S., & Coleman, P. K. (2001). Deception, power, and self-differentiation in college students' romantic relationships: An exploratory study. *Journal of Sex & Material Therapy*, *27*, 57–71.

Blandon-Gitlin, I., Pezdek, K., Rogers, M., & Brodie, L. (2005). Detecting deception in children: An experimental study of the effect of event familiarity on CBCA ratings. *Law and Human Behavior*, *29*, 187–197.

Bleske. A. L., & Shackelford, T. K. (2001). Poaching, promiscuity, and deceit: Combatting mating rivalry in same-sex friendships. *Personal Relationships*, *8*, 407–424.

Boaz, T. L., Perry, N. W., Raney, G., Fischler, I. S., & Shuman, D. (1991). Detection of guilty knowledge with event-related potentials. *Journal of Applied Psychology*, *76*, 788–795.

Bok, S. (1978). *Lying: Moral choice in public and private life*. New York: Random House.

Boltz, M. G. (2005). Temporal dimensions of conversational interaction: The role of response latencies and pauses in social impression formation. *Journal of Language and Social Psychology*, *24*, 103–138.

Bond, C. F. (in press). A few can catch a liar, sometimes. *Applied Cognitive Psychology*.

Bond, C. F., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin*, *26*, 385–395.

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

Bond, C. F., & DePaulo, B. M. (2007). *Individual differences in detecting deception*. Manuscript submitted for publication.

Bond, C. F., & Fahey, W. E. (1987). False suspicion and the misperception of deceit. *British Journal of Social Psychology*, *26*, 41–46.

Bond, C. F., Kahler, K. N., & Paolicelli, L. M. (1985). The miscommunication of deception: An adaptive perspective. *Journal of Experimental Social Psychology*, *21*, 331–345.

Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior*, *14*, 189–205.

Bond, C. F., Omar, A., Pitre, U., Lashley, B. R., Skaggs, L. M. & Kirk, C. T. (1992). Fishy-looking liars: Deception judgment from expectancy violation. *Journal of Personality and Social Psychology*, *63*, 969–977.

Bond, C. F., & Rao, S. R. (2004). Lies travel: Mendacity in a mobile world. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 127–147). Cambridge, England: Cambridge University Press.

Bond, C. F., & Robinson, M. (1988). The evolution of deception. *Journal of Nonverbal Behavior*, *12*, 295–307.

Bond, C. F., Thomas, B. J., & Paulson, R. M. (2004). Maintaining lies: The multiple-audience problem. *Journal of Experimental Social Psychology*, *40*, 29–40.

Bond, C. F., & Uysal, A. (2007). On lie detection 'wizards'. *Law and Human Behavior*, *31*, 109–115.

Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology*, *19*, 313–329.

Bond, G. D., Malloy, D. M., Arias, E. A., & Nunn, S., & Thompson, L. A (2005). Lie-biased decision making in prison. *Communication Reports*, *18*, 9–19.

Bond, G. D., Malloy, D. M., Thompson, L. A., Arias, E. A., & Nunn, S. (2004). Post-probe decision making in a prison context. *Communication Monographs*, *71*, 269–285.

Book, A. S., Holden, R. R., Starzyk, K. B., Wasylkiw, L., & Edwards, M. J. (2006). Psychopathic traits and experimentally induced deception in self-report assessment. *Personality and Individual Differences*, *41*, 601–608.

Boon, S. D., & McLeod, B. A. (2001). Deception in romantic relationships: Subjective estimates of success at deceiving and attitudes toward deception. *Journal of Social and Personal Relationships*, *18*, 463–476.

Bothwell, R., & Jalil, M. (1992). The credibility of nervous witnesses. *Journal of Social Behavior and Personality*, *7*, 581–586.

Bowlby, J. (1969). *Attachment and loss: Vol. 1 attachment*. New York: Basic Books.

Bowlby, J. (1973). *Attachment and loss: Vol. 2 separation: anxiety and anger*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and loss: Vol. 3 loss*. New York: Basic Books.

Boychuk, T. (1991). *Criteria-Based Content Analysis of children's statements about sexual abuse: A field-based validation study*. Unpublished doctoral dissertation, Arizona State University.

Bradford, R. (1994). Developing an objective approach to assessing allegations of sexual abuse. *Child Abuse Review*, *3*, 93–101.

Bradley, M. T., & Janisse, M. P. (1981). Accuracy demonstrations, threat, and the detection of deception: Cardiovascular electrodermal, and pupillary measures. *Psychophysiology*, *18*, 307–315.

Branaman, T. F., & Galagher, S. N. (2005). Polygraph testing in sex offender treatment: A review of limitations. *American Journal of Forensic Psychology*, *23*, 45–64.

Brandt, D. R., Miller, G. R., & Hocking, J. E. (1980a). The truth-deception attribution: Effects of familiarity on the ability of observers to detect deception. *Human Communication Research*, *6*, 99–110.

Brandt, D. R., Miller, G. R., & Hocking, J. E. (1980b). Effects of self-monitoring and familiarity on decepion detection. *Communication Quarterly*, *28*, 3–10.

Brandt, D. R., Miller, G. R., & Hocking, J. E. (1982). Familiarity and lie detection: A replication and extension. *The Western Journal of Speech Communication*, *46*, 276–290.

Brennan, K., A., Clark, C. L., & Shaver, P. R. (1998). Self-report measures of adult attachment: An integrative overview. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 46–76). New York: Guilford Press.

Brigham, J. C. (1999). What is forensic psychology, anyway? *Law and Human Behavior*, *23*, 273–298.

British Psychological Society (1986). Report of the Working Group on the use on the polygraph in criminal investigation and personnel screening. *Bulletin of the British Psychological Society*, *39*, 81–94.

British Psychological Society (2004). *A review of the current scientific status and fields of application of polygraphic deception detection. Final report from the BPS Working Party*. Leicester, England: British Psychological Society.

Brooks, C. I., Church, M. A., & Fraser, L. (1986). Effects of duration of eye contact on judgments of personality characteristics. *The Journal of Social Psychology*, *126*, 71–78.

Broomfield, K. A., Robinson, E. J., & Robinson, W. P. (2002). Children's understanding about white lies. *British Journal of Developmental Psychology*, *20*, 47–65.

Brougham, C. G. (1992). Nonverbal communication: Can what they don't say give them away? *FBI Law Enforcement Bulletin*, *61*, 15–18.

Bruck, M., Ceci, S. J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review*, *22*, 520–554.

Buck, J. A., Warren, A. R., Betman, S., & Brigham, J. C. (2002). Age differences in Criteria-Based Content Analysis scores in typical child sexual abuse interviews. *Applied Developmental Psychology, 23*, 267–283.

Bugental, D. B., Shennum, W., Frank, M., & Ekman, P. (2001). "True lies": Children's abuse history and power attributions as influences on deception detection. In V. Manusov & J. H. Harvey (Eds.), *Attribution, communication behavior, and close relationships* (pp. 248–265). Cambridge, England: Cambridge University Press.

Bull, R. (1988). What is the lie-detection test? In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 10–19). London: Sage.

Bull, R. (1992). Obtaining evidence expertly: The reliability of interviews with child witnesses. *Expert Evidence: The International Digest of Human Behaviour Science and Law*, *1*, 3–36.

Bull, R. (1995). Innovative techniques for the questioning of child witnesses, especially those who are young and those with learning disability. In M. Zaragoza *et al.* (Eds.), *Memory and testimony in the child witness* (pp. 179–195). Thousand Oaks, CA: Sage.

Bull, R. (2004). Training to detect deception from behavioural cues: Attempts and problems. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 251–268) Cambridge, England: Cambridge University Press.

Bull, R. (2006). *The views of the British Pstychological Society Working Party on scientific status and fields of application of polygraph testing*. Paper presented at the European Expert Meeting on Polygraph Testing. University of Maastricht, Maastricht, the Netherlands, 29–31 March.

Bull, R., & Rumsey, N. (1988). *The social psychology of facial appearance*. New York: Springer-Verlag.

Buller, D. B., & Aune, R. K. (1987). Nonverbal cues to deception among intimates, friends, and strangers. *Journal of Nonverbal Behavior*, *11*, 269–289.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, *6*, 203–242.

Buller, D. B., Burgoon, J. K., Busling, A. L. & Roiger, J. F. (1994a). Interpersonal deception: VIII. Further analysis of nonverbal and verbal correlates of equivocation from the Bauelas et al. (1990) research. *Journal of Language and Social Psychology*, *13*, 396–417.

Buller, D. B., Burgoon, J. K., White, C. H., & Ebesu, A. S. (1994b). Interpersonal deception: VII. Behavioral profiles of falsification, equivocation and concealment. *Journal of Language and Social Psychology*, *13*, 366–395.

Buller, D. B., Comstock, J., Aune, R. K., & Strzyzewski, K. D. (1989). The effect of probing on deceivers and truthtellers. *Journal of Nonverbal Behavior*, *13*, 155–170.

Buller, D. B., Stiff, J. B., & Burgoon, J. K. (1996). Behavioral adaptation in deceptive transactions. Fact or fiction: A reply to Levine and McCornack. *Human Communication Research*, *22*, 589–603.

Buller, D. B., Strzyzewski, K. D., & Comstock, J. (1991). Interpersonal deception: I. Deceivers' reactions to receivers' suspicions and probing. *Communication Monographs*, *58*, 1–24.

Buller, D. B., Strzyzewski, K. D., & Hunsaker, F. G. (1991). Interpersonal deception: II. The inferiority of conversational participants as deception detectors. *Communication Monographs*, *58*, 25–40.

Burgess, A. W. (1985). *Rape and sexual assault: A research book*. London: Garland.

Burgess, A. W., & Homstrom, L. L. (1974). *Rape: Victims of crisis*. Bowie: Brady.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal dynamics. *Journal of Nonverbal Behavior*, *18*, 155–184.

Burgoon, J. K., Buller, D. B., Dillman, L., & Walther, J. B. (1995). Interpersonal deception IV: Effects of suspicion on perceived communication and nonverbal behaviour dynamics. *Human Communication Research*, *22*, 163–196.

Burgoon, J. K., Buller, D. B., Ebesu, A. S., White, C. H., & Rockwell, P. A. (1996a). Testing interpersonal deception theory: Effects of suspicion on communication behaviors and perception. *Communication Theory*, *6*, 243–267.

Burgoon, J. K., Buller, D. B., & Floyd, K. (2001). Does participation affect deception success? A test of the interactivity principle. *Human Communication Research*, *27*, 503–534.

Burgoon, J. K., Buller, D. B., Floyd, K., & Grandpre, J. (1996). Deceptive realities: Sender, receiver, and observer perspectives in deceptive conversations. *Communication Research*, *23*, 724–748.

Burgoon, J. K., Buller, D. B., Guerrero, L. K., Afifi, W. A., & Feldman, C. M. (1996b). Interpersonal deception: XII. Information management dimensions underlying deceptive and truthful messages. *Communication Monographs*, *63*, 50–69.

Burgoon, J. K., Buller, D. B., White, C. H., Afifi, W., & Buslig, A. L. S. (1999). The role of conversation involvement in deceptive interpersonal interactions. *Personality and Social Psychology Bulletin*, *25*, 669–685.

Burgoon, J. K., & Qin, T. (2006). The dynamic nature of deceptive verbal communication. *Journal of Language and Social Psychology*, *25*, 76–96.

Burns, J. A., & Kintz, B. L. (1976). Eye contact while lying during an interview. *Bulletin of the Psychonomic Society*, *7*, 87–89.

Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, *50*, 559–570.

Bussey, K. (1992). Children's lying and truthfulness: Implications for children's testimony. In S. J. Ceci, M. DeSimone Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 89–110). Hillsdale, NJ: Erlbaum.

Bybee, D., & Mowbray, C. T. (1993). An analysis of allegations of sexual abuse in a multi-victim day-care center case. *Child Abuse and Neglect*, *17*, 767–783.

Cappella, J. N., & Schreiber, D. M. (2006). The interaction management function of nonverbal cues. In V. Manusov & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 361–379). Thousand Oaks, CA: Sage.

Carmel, D., Dayan, E., Naveh, A., Raveh, O., & Ben-Shakhar, G. (2003). Estimating the validity of the guilty knowledge test from simulated experiments: The external validity of mock crime studies. *Journal of Experimental Psychology: Applied*, *9*, 261–269.

Carpenter, R. H. (1981). Stylistic analysis for law enforcement purposes: A case study of a language variable as an index of a suspect's caution in phrasing answers. *Communication Quarterly*, *29*, 32–39.

Carroll, D. (1988). How accurate is polygraph lie detection? In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 20–28). London: Sage.

Carroll, D. (1991). Lie detection: Lies and truths. In R. Cochrane, & D. Carroll (Eds.), *Psychology and social issues: A tutorial test* (pp. 160–170). London: Falmer Press.

Caso, L., Gnisci, A., Vrij, A., & Mann, S. (2005). Processes underlying deception: An empirical analysis of truths and lies when manipulating the stakes. *Journal of Interviewing and Offender Profiling*, *2*, 195–202.

Caso, L., Maricchiolo, F., Bonaiuto, M., Vrij, A., & Mann, S. (2006a). The impact of deception and suspicion on different hand movements. *Journal of Nonverbal Behavior*, *30*, 1–19.

Caso, L., Vrij, A., Mann, S., & DeLeo, G. (2006b). Deceptive responses: The impact of verbal and nonverbal countermeasures. *Legal and Criminological Psychology*, *11*, 99–111.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom*. Washington, DC: American Psychological Association.

Ceci, S. J., & Bruck, M. (1998). Reliability and credibility of young children's reports. *American Psychologist*, *53*, 136–151.

Ceci, S. J., Crossman, A. M., Scullin, M. H., Gilstrap, L., & Huffman, M. L. (2002). Children's suggestibitilty research: Implications for the courtroom and the forensic interview. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 117–130). Chichester, England: John Wiley & Sons, Ltd.

Ceci, S. J., & DeSimone Leichtman, M. (1992). "I know that you know that I know that you broke the toy": A brief report of recursive awareness among 3-year-olds. In S. J. Ceci, M. DeSimone Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 1–9). Hillsdale, NJ: Erlbaum.

Ceci, S. J., Huffman, M. L., Smith, E., & Loftus, E. F. (1994). Repeatedly thinking about a non-event. *Consciousness and Cognition*, *3*, 388–407.

Ceci, S. J., Loftus, E. F., Leichtman, M. D., & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis*, *17*, 304–320.

Chahal, K., & Cassidy, T. (1995). Deception and its detection in children: A study of adult accuracy. *Psychology, Crime, & Law*, *1*, 237–245.

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development*, *60*, 1263–1277.

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior*, *6*, 151–155.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Cherry, E. C. (1953) Some experiments on the recognition of speech, with one and with two ears. *Journal of Acoustic Society of America, 25*, 975–979.

Chiba, H. (1985). Analysis of controlling facial expression when experiencing negative affect on an anatomical basis. *Journal of Human Development, 21*, 22–29.

Christianson, S. A. (Ed.), (2007). *Offenders' memories of violent crimes*. Chichester, England: John Wiley & Sons, Ltd.

Cody, M. J., Lee, W. S., & Chao, E. Y. (1989). Telling lies: Correlates of deception among Chinese. In J. P. Forgas & J. M. Innes (Eds.), *Recent advances in social psychology: An international perspective* (pp. 359–368). North Holland, the Netherlands: Elsevier.

Cody, M. J., Marston, P. J., & Foster, M. (1984). Deception: Paralinguistic and verbal leakage. In R. N. Bostrom, & B. H. Westley (Eds.), *Communication Yearbook 8* (pp. 464–490). Beverly Hills, CA: Sage.

Cody, M. J., & O'Hair, H. D. (1983). Nonverbal communication and deception: Differences in deception cues due to gender and communicator dominance. *Communication Monographs*, *50*, 175–193.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cole, T. (2001). Lying to the one you love: The use of deception in romantic relationships. *Journal of Social and Personal Relationships*, *18*, 107–129.

Cole, T., Leets, L., & Bradac, J. J. (2002). Deceptive message processing: The role of attachment style and verbal intimacy markers in deceptive message judgments. *Communication Studies*, *53*, 74–89.

Colwell, K., Hiscock, C. K., & Memon, A. (2002). Interview techniques and the assessment of statement credibility. *Applied Cognitive Psychology*, *16*, 287–300.

Colwell, K., Hiscock-Anisman, C., Memon, A., Woods, D., & Michlik, P. M. (2006). Strategies of impression management among deceivers and truth tellers: How liars attempt to convince. *Amercian Journal of Forensic Psychology*, *24*, 31–38.

Colwell, L. H., Miller, H. A., Lyons, P. M., & Miller, R. S. (2006). The training of law enforcement officers in detecting deception: A survey of current practices and suggestions for improving accuracy. *Police Quarterly*, *9*, 275–290.

Colwell, L. H., Miller, H. A., Miller, R. S., & Lyons, P. M. (2006). US police officers' knowledge regarding behaviors indicative of deception: Implications for eradicating erroneous beliefs through training. *Psychology, Crime, & Law*, *12*, 489–503.

Comadena, M. E. (1982). Accuracy in detecting deception: Intimate and friendship relationships. In M. Burgoon (Ed.), *Communication yearbook 6* (pp. 446–472). Beverly Hills, CA: Sage.

Consigli, J. E. (2002). Post-conviction sex offender testing and the American polygraph association. In M. Kleiner (Ed.). *Handbook of polygraph testing* (pp. 237–250). San Diego, CA: Academic Press.

Conte, J. R., Sorenson, E., Fogarty, L., & Rosa, J. D. (1991). Evaluating children's reports of sexual abuse: Results from a survey of professionals. *Journal of Orthopsychiatry*, *61*, 428–437.

Coolbear, J. L. (1992). Credibility of young children in sexual abuse cases: Assessment strategies of legal and human service professionals. *Canadian Psychology*, *33*, 151–164.

Cordon, I. M., Pipe, M. E., Sayfan, L., Melinder, A., & Goodman, G. S. (2004). Memory for traumatic experiences in childhood. *Developmental Review*, *24*, 101–132.

Craig, R. A. (1995). *Effects of interviewer behavior on children's statements of sexual abuse*. Unpublished manuscript.

Craig, R. A., Scheibe, R., Raskin, D. C., Kircher, J. C., & Dodd, D. H. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. *Applied Developmental Science*, *3*, 77–85.

Craig, K. D., Hyde, S. A., & Patrick, C. J. (1991). Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain. *Pain*, *46*, 161–171.

Crombag, H. F. M., Wagenaar, W. A., & Van Koppen, P. J. (1996). Crashing memories and the problem of source monitoring. *Applied Cognitive Psychology*, *10*, 93–104.

Cross, T. P., & Saxe, L. (2001). Polygraph testing and sexual abuse: The lure of the magic lasso. *Child Maltreatment*, *6*, 195–206.

Crossman, A. M., & Lewis, M. (2006). Adults' ability to detect children's lying. *Behavioral Sciences and the Law*, *24*, 703–715.

Cullen, M. C., & Bradley, M. T. (2004). Positions of truthfully answered controls on Control Question Tests with the polygraph. *Canadian Journal of Behavioural Science*, *36*, 167–176.

Curtis, R. C., & Miller, K. (1986). Believing another likes or dislikes you: Behaviors making the beliefs come true. *Journal of Personality and Social Psychology*, *51*, 284–290.

Cutler, B. L., Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behavior*, *14*, 185–191.

Cutler, B. L., Penrod, S. D., & Stuve, T. E. (1988). Juror decision making in eyewitness identification cases. *Law and Human Behavior*, *12*, 41–55.

Dalenberg, C. J., Hyland, K. Z., & Cuevas, C. A. (2002). Sources of fantastic elements in allegations of abuse by adults and children. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 185–204). Mahwah, NJ: Erlbaum.

Dalton, A., L., & Daneman, M. (2006). Social suggestibility to central and peripheral misinformation. *Memory*, *14*, 486–501.

Daniels. C. W. (2002). Legal aspects of polygraph admissibility in the United Stated. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 327–338). San Diego, CA: Academic Press.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labelling effects. *Journal of Personality and Social Psychology*, *44*, 20–33.

Darling, N., Cumsille, P., Caldwell, L. L., & Dowdy, B. (2006). Predictors of adolescents's disclosure to parents and perceived parental knowledge: Between- and within-person differerences. *Journal of Youth and Adolescence*, *35*, 667–678.

Daubert v Merrell Dow Pharmacurticals, Inc. 113 S. Ct. 2786, 1993.

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., Gur, R. C., & Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine learning methods: Application to lie detection. *NeuroImage*, *28*, 663–668.

Davies, G. M. (1991). Research on children's testimony: Implications for interviewing practice. In C. R. Hollin & K. Howells (Eds.), *Clinical approaches to sex offenders and their victims* (pp. 177–191). New York: John Wiley & Sons, Inc.

Davies, G. M. (1994a). Children's testimony: Research findings and police implications. *Psychology, Crime, & Law*, *1*, 175–180.

Davies, G. M. (1994b). Statement validity analysis: An art or a science? Commentary on Bradford. *Child Abuse Review*, *3*, 104–106.

Davies, G. M. (2001). Is it possible to discriminate true from false memories? In G. M. Davies & T. Dalgleish (Eds.), *Recovered memories: Seeking the middle ground* (pp. 153–176). Chichester, England: John Wiley & Sons, Ltd.

Davies, G. M. (2004). Coping with suggestion and deception in children's accounts. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 148–171). Cambridge, England: Cambridge University Press.

Davies, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime, & Law*, *6*, 81–97.

Davis, M., & Hadiks, D. (1995). Demeanor and credibility. *Semiotica*, *106*, 5–54.

Davis, M., Markus, K. A., & Walters, S. B. (2006). Judging the credibility of criminal suspect statements: Does mode of presentation matter? *Journal of Nonverbal Behavior*, *30*, 181–198.

Davis, M., Markus, K. A., Walters, S. B., Vorus, N., & Connors, B. (2005). Behavioral cues to deception vs topic incriminating potential in criminal confessions. *Law and Human Behavior*, *29*, 683–704.

Davis, R. C. (1961). Psychological responses as a means of evaluating information. In A. Biderman & H. Zimmer (Eds.), *Manipulation of human behavior* (pp. 142–168). New York: John Wiley & Sons, Inc.

De Keijser, J. W., & van Koppen, P. J. (2007). Paradoxes of proof and punishment: Psychological pitfalls in judicial decision-making. *Legal and Criminological Psychology*, *12*, 189–206.

Dente, E., Barath, A. A., Ng, J., Vrij, A., Mann, S., & Bull, A. (2006). Tracking hand and finger movements for behaviour analysis. *Pattern Recognition Letter*, *27*, 1797–1808.

DePaulo, B. M. (1988). Nonverbal aspects of deception. *Journal of Nonverbal Behavior*, *12*, 153–162.

DePaulo, B. M. (1991). Nonverbal behavior and self-presentation: A developmental perspective. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 351–397). Paris, France: Cambridge University Press.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, *111*, 203–243.

DePaulo, B. M. (1994). Spotting lies: Can humans learn to do better? *Current Directions in Psychological Science*, *3*, 83–86.

DePaulo, B. M. (2004). The many faces of lies. In A. G. Miller (Ed.), *The social psychology of good and evil* (pp. 303–336). New York, Guilford Press.

DePaulo, B. M., Ansfield, M. E., & Bell, K. L. (1996). Theories about deception and paradigms for studying it: A critical appraisal of Buller and Burgoon's interpersonal deception theory and research. *Communication Theory*, 297–310.

DePaulo, B. M., Ansfield, M. E., Kirkendol, S. E., & Boden, J. M. (2004). Serious lies. *Basic and Applied Social Psychology*, *26*, 147–167.

DePaulo, B. M., & Bell, K. L. (1996). Truth and investment: Lies are told to those who care. *Journal of Personality and Social Psychology*, *70*, 703–716.

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. L., & Muhlenbruck, L. (1997). The accuracy – confidence correlation in the detection of deception. *Personality and Social Psychology Review*, *1*, 346–357.

DePaulo, B. M., Epstein, J. A., & LeMay, C. S. (1990). Responses of the socially anxious to the prospect of interpersonal evaluation. *Journal of Personality*, *58*, 623–640.

DePaulo, B. M., Epstein, J. A., & Wyer, M. M. (1993). Sex differences in lying: How women and men deal with the dilemma of deceit. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 126–147). New York: Guilford Press.

DePaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 3–40). Boston, MA: McGraw-Hill.

DePaulo, B. M., & Jordan, A. (1982). Age changes in deceiving and detecting deceit. In R. S. Feldman (Ed.), *Development of nonverbal behaviour in children* (pp. 151–180). New York: Springer-Verlag.

DePaulo, B. M., Jordan, A., Irvine, A., & Laser, P. S. (1982). Age changes in the detection of deception. *Child Development*, *53*, 701–709.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology*, *74*, 63–79.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*, 979–995.

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., Kirkendol, S. E., Tang, J., & O'Brien, T. P. (1988). The motivational impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior*, *12*, 177–202.

DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting deceit of the motivated liar. *Journal of Personality and Social Psychology*, *45*, 1096–1103.

DePaulo, B. M., Lassiter, G. D., & Stone, J. I. (1982a). Attentional determinants of success at detecting deception and truth. *Personality and Social Psychology Bulletin*, *8*, 273–279.

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.

DePaulo, B. M., & Morris, W. L. (2004). Discerning lies from truths: Behavioural cues to deception and the indirect pathway of intuition. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 15–40). Cambridge, England: Cambridge University Press.

DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, *16*, 249–267.

DePaulo, B. M., Rosenthal, R., Eisenstat, R. A., Rogers, P. L., & Finkelstein, S. (1978). Decoding discrepant nonverbal cues. *Journal of Personality and Social Psychology*, *36*, 313–323.

DePaulo, B. M., Rosenthal, R., Rosenkrantz, & Green, C. R. (1982b). Actual and perceived cues to deception: A closer look at speech. *Basic and Applied Social Psychology*, *3*, 291–312.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Telling ingratiating lies: Effects of target sex and target attractiveness on verbal and nonverbal deceptive success. *Journal of Personality and Social Psychology*, *48*, 1191–1203.

DePaulo, B. M., & Tang, J. (1994). Social anxiety and social judgement: The example of detecting deception. *Journal of Research in Personality*, *28*, 142–153.

DePaulo, B. M., Tang, J., & Stone, J. L. (1987). Physical attractiveness and skill at detecting deception. *Personality and Social Psychology Bulletin*, *13*, 177–187.

DePaulo, B. M., Wetzel, C., Weylin Sternglanz, R., & Walker Wilson, M. J. (2003). Verbal and nonverbal dynamics of privacy, secrecy, and deceit. *Journal of Social Issues*, *59*, 391–410.

DePaulo, P. J., & DePaulo, B. M. (1989). Can deception by salespersons and customers be detected through nonverbal behavioral cues? *Journal of Applied Social Psychology*, *19*, 1552–1577.

Desforges, D. M., & Lee, T. C. (1995). Detecting deception is not as easy as it looks. *Teaching of Psychology*, *22*, 128–130.

deTurck, M. A. (1991). Training observers to detect spontaneous deception: Effects of gender. *Communication Reports*, *4*, 81–89.

deTurck, M. A., Feeley, T. H., & Roman, L. A. (1997). Vocal and visual cue training in behavioural lie detection. *Communication Research Reports*, *14*, 249–259.

deTurck, M. A., Harszlak, J. J., Bodhorn, D. J., & Texter, L. A. (1990). The effects of training social perceivers to detect deception from behavioural cues. *Communication Research*, *38*, 189–199.

deTurck, M. A., & Miller, G. R. (1985). Deception and arousal: Isolating the behavioral correlates of deception. *Human Communication Research*, *12*, 181–201.

deTurck, M. A., & Miller, G. R. (1990). Training observers to detect deception: Effects of self-monitoring and rehearsal. *Human Communication Research*, *16*, 603–620.

DeWaal, F. (1986). Deception in the natural communication of chimpanzees. In R. W. Mitchell & N. S. Thompson (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 221–244). Albany: State University of New York Press.

Dimberg, U., Thunberg, M., & Grunedal, S. (2002). Facial reactions to emotional stimuli: Automatic controlled emotional responses. *Cognition and Emotion*, *16*, 449–471.

Doherty-Sneddon, G., Bruce, V., Bonner, L., Longbotham, S., & Doyle, C. (2002). Development of gaze aversion as disengagement of visual information. *Developmental Psychology*, *38*, 438–445.

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory and Cognition*, *33*, 727–733.

Donaghy, W. C., & Dooley, B. F. (1994). Head movement, gender, and deceptive communication. *Communication Reports*, *7*, 67–75.

Doris, J. (1994). Commentary on Criteria-Based Content Analysis. *Journal of Applied Developmental Psychology*, *15*, 281–285.

Driscoll, L. N. (1994). A validity assessment of written statements from suspects in criminal investigations using the SCAN technique. *Police Studies*, *17*, 77–88.

Dulaney, E. F. (1982). Changes in language behavior as a function of veracity. *Human Communication Research*, *9*, 75–82.

Durban, N. E., Ramirez, A., & Burgoon, J. K. (2003). The effects of participation on the ability to judge deceit. *Communication Reports*, *16*, 23–33.

Ebesu, A. S., & Miller, M. D. (1994). Verbal and nonverbal behaviors as a function of deception type. *Journal of Language and Social Psychology*, *13*, 418–442.

Edelstein, R. S., Luten, T. L., Ekman, P., & Goodman, G. S. (2006). Detecting lies in children and adults. *Law and Human Behavior*, *30*, 1–10.

Edinger, J. A., & Patterson, M. L. (1983). Nonverbal involvement and social control. *Psychological Bulletin*, *93*, 30–56.

Efron, D. (1941). *Gesture and environment*. New York: King's Crown.

Eichenbaum, H., & Bodkin, J. A. (2000). Belief and knowledge as distinct forms of memory. In D. L. Schacter & E. Scaryy (Eds.), *Memory, brain, and belief* (pp. 176–207). Cambridge, MA: Harvard University Press.

Ekman, P. (1981). Mistakes when deceiving. *Annals of the New York Academy of Sciences*, *364*, 269–278.

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, politics and marriage*. New York: W. W. Norton. (Reprinted in 1992 and 2001).

Ekman, P. (1988). Lying and nonverbal behavior: Theoretical issues and new findings. *Journal of Nonverbal Behavior*, *12*, 163–176.

Ekman, P. (1989). Why lies fail and what behaviors betray a lie. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 71–82). Dordrecht, the Netherlands: Kluwer.

Ekman, P. (1993). Why don't we catch liars? *Social Research*, *63*, 801–817.

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: Oxford University Press.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology II. *Journal of Personality and Social Psychology*, *58*, 342–353.

Ekman, P., & Frank, M. G. (1993). Lies that fail. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 184–201). New York: Guilford Press.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*, 88–106.

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

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, *29*, 288–298.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior*, *6*, 238–253.

Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology*, *54*, 414–420.

Ekman, P., Friesen, W. V., & O'Sullivan, M., & Scherer, K. (1980). Relative importance of face, body and speech in judgements of personality and affect. *Journal of Personality and Social Psychology*, *38*, 270–277.

Ekman, P., Friesen, W. V., & Scherer, K. R. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica*, *16*, 23–27.

Ekman, P., Friesen, W. V., & Simons, R. C. (1985). Is the startle reaction an emotion? *Journal of Personality and Social Psychology*, *49*, 1416–1426.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, 46, 913–920.

Ekman, P., & O'Sullivan, M. (2006). From flawed self-assessment to blatant whoppers: The utility of voluntary and involuntary behavior in detecting deception. *Behavioural Sciences & the Law*, *24*, 673–686.

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science*, *10*, 263–266.

Ekman, P., O'Sullivan, M., Friesen, W. V., & Scherer, K. (1991). Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior*, *15*, 125–135.

Ekman, P., Roper, G., & Hager, J. C. (1980). Deliberate facial movement. *Child Development*, *51*, 886–891.

Elaad, E. (1990). Detection of guilty knowledge in real-life criminal investigations. *Journal of Applied Psychololgy*, *75*, 521–529.

Elaad, E. (1993). Detection of deception: A transactional analysis perspective. *Journal of Psychology*, *127*, 5–15.

Elaad, E. (1997). Polygraph examiner awareness of crime-relevant information and the guilty knowledge test. *Law and Human Behavior*, *21*, 107–120.

Elaad, E. (2003). Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology*, *17*, 349–363.

Elaad, E., & Ben-Shakhar, G. (1997). Effects of item repetitions and variations on the efficiency of the guilty knowledge test. *Psychophysiology*, *34*, 587–596.

Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real-life criminal guilty knowledge tests. *Journal of Applied Psychology*, *77*, 757–767.

Elaad, E., Ginton, A., & Ben-Shakhar, G. (1994). The effects of prior expectations and outcome knowledge on polygraph examiners' decisions. *Journal of Behavioral Decision Making*, *7*, 279–292.

Ellwanger, J., Rosenfeld, J. P., Hannkin, L. B., & Sweet, J. J. (1999). P300 as an index of recognition in a standard and difficult to match-to-sample test: A model of amnesia in normal adults. *The Clinical Neuropsychologist*, *13*, 100–108.

Ellwanger, J., Rosenfeld, J. P., Sweet, J. J., & Bhatt, M. (1996). Detecting dimulated amnesia for autobiographical and recently learned information using the P300 event-related potential. *International Journal of Psychophysiology*, *23*, 9–23.

Engelhard, I. M., Merckelbach, H., & van den Hout, M. A. (2003). The guilty knowledge test and the modified stroop task in detection of deception: An exploratory study. *Psychological Reports*, *92*, 683–691.

Engels, R. C. M. E., Finkenauer, C., & van Kooten, D. C. (2006). Lying behaviour, family functioning and adjustment in early adolescence. *Journal of Youth Adolescence*, *35*, 949–958.

Ennis, E., Vrij, A., & Chance, C. (in press). Individual differences and lying in everyday life. *Journal of Social and Personal Relationships*.

Erdmann, K., Volbert, R., & Böhm, C. (2004). Children report suggested events even when interviewed in a non-suggestive manner: What are the implications for credibility assessment? *Applied Cognitive Psychology*, *18*, 589–611.

Esplin, P. W., Boychuk, T., & Raskin, D. C. (1988, June). *A field validity study of Criteria-Based Content Analysis of children's statements in sexual abuse cases*. Paper presented at the NATO Advanced Study Institute on Credibility Assessment in Maratea, Italy.

Etcoff, N. L., Ekman, P., Magee, J. J., & Frank, M. G. (2000). Lie detection and language comprehension. *Nature*, *405*, 139.

Evans, R. (1994). Police interviews with juveniles. In D. Morgan & G. M. Stephenson (Eds.), *Suspicion and silence: The right to silence in criminal investigations* (pp. 77–90). London: Blackstone.

Exline, R., Thibaut, J., Hickey, C., & Gumpert, P. (1970). Visual interaction in relation to Machiavellianism and an unethical act. In P. Christie & F. Geis (Eds.), *Studies in Machiavellianism* (pp. 53–75). New York: Academic Press.

Eyre, S. L., Read, N. W., & Millstein, S. G. (1997). Adolescent sexual strategies. *Journal of Adolescent Health*, *20*, 286–293.

Eysenck, H. J. (1984). Crime and personality. In D. J. Muller, D. E. Blackman, & A. J. Chapman (Eds.), *Psychology and Law* (pp. 85–100). New York: John Wiley & Sons, Inc.

Faigman, D. L., Kaye, D., Saks, M. J., & Sanders, J. (1997). *Modern scientific evidence: The law and science of expert testimony*. St. Paul, MN: West Publishing.

Faigman, D. L., Kaye, D., Saks, M. J., & Sanders, J. (2002). *Modern scientific evidence: The law and science of expert testimony, volume 2*. St. Paul, MN: West Publishing.

Farrow, T. F. D., Reilly, R., Rahman, T. A., Herford, A. E., Woodruff, P. W. R., & Spence, S. A. (2003). Sex and personality traits influence the difference between time taken to tell the truth or lie. *Perceptual and Motor Skills*, *97*, 451–460.

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ('lie detection') with event-related brain potentials. *Psychophysiology*, *28*, 531–547.

Farwell, L. A., & Smith, S. S. (2001). Using brain MERMER testing to detect knowledge despite efforts to conceal. *Journal of Forensic Sciences*, *46*, 135–143.

Feeley, T. H., & deTurck, M. A. (1995). Global cue usage in behavioural lie detection. *Communication Quarterly*, *43*, 420–430.

Feeley, T. H., & deTurck, M. A. (1997). *Perceptions of communication as seen by the actor and as seen by the observer: The case of lie detection*. Paper presented at the International Communication Association Annual Conference. Montreal, Canada.

Feeley, T. H., & deTurck, M. A. (1998). The behavioral correlates of sanctioned and unsanctioned deceptive communication. *Journal of Nonverbal Behavior*, *22*, 189–204.

Feeley, T. H., deTurck, M. A., & Young, M. J. (1995). Baseline familiarity in lie detection. *Communication Research Reports*, *12*, 160–169.

Feeley, T. H., & Young, M. J. (2000). The effects of cognitive capacity on beliefs about deceptive communication. *Communication Quarterly*, *48*, 101–119.

Feldman, R. S. (1979). Nonverbal disclosure of deception in urban Koreans. *Journal of Cross-Cultural Psychology*, *10*, 73–83.

Feldman, R. S., & Chesley, R. B. (1984). Who is lying, who is not: An attributional analysis of the effects of nonverbal behavior on judgments of defendant believability. *Behavioral Sciences and the Law*, *2*, 451–461.

Feldman, R. S., Devin-Sheehan, L., & Allen, V. L. (1978). Nonverbal cues as indicators of verbal dissembling. *American Educational Research Journal*, *15*, 217–231.

Feldman, R. S., Forrest, J. A., & Happ, B. R. (2002). Self-presentation and verbal deception: Do self-presenters lie more? *Basic and Applied Social Psychology*, *24*, 163–170.

Feldman, R. S., Jenkins, L., & Popoola, O. (1979). Detection of deception in adults and children via facial expressions. *Child Development*, *50*, 350–355.
Feldman, R. S. & Phillippot, P. (1993). Children's deception skills and social competence. In G. Goodman, & B. Bottoms (Eds.), *Child victims, child witnesses* (p. 80–99). New York: Guilford Press.
Feldman, R. S., Tomasian, J. C., & Coats, E. J. (1999). Nonverbal deception abilities and adolescents' social competence: Adolescents with higher social skills are better liars. *Journal of Nonverbal Behavior*, *23*, 237–250.
Feldman, R. S., & White, J. B. (1980). Detecting deception in children. *Journal of Communication*, 121–128.
Feldman, R. S., White, J. B., & Lobato, D. (1982). Social skills and nonverbal behavior. In R. S. Feldman (Ed.), *Development of nonverbal behavior in children*. New York: Springer-Verlag.
Feldman, S. S., & Cauffman, E. (1999). Sexual betrayal among late adolescents: Perspectives of the perpetrator and the aggrieved. *Journal of Youth and Adolescence*, *28*, 235–258.
Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*, 522–527.
Fiedler, K., Schmid, J., & Stahl, T. (2002). What is the current truth about polygraph lie detection? *Basic and Applied Social Psychology*, *24*, 313–324.
Fiedler, K., & Walka, I. (1993). Training lie detectors to use nonverbal cues instead of global heuristics. *Human Communication Research*, *20*, 199–223.
Fisher, R. P., Brennan, K. H., & McCauley, M. R. (2002). The cognitive interview method to enhance eyewitness recall. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 265–286). Mahwah, NJ: Erlbaum.
Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The cognitive interview*. Springfiled, IL: Charles C. Thomas.
Fivush, R. (2002). The development of autobiograhical memory. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 55-68). Chichester, England: John Wiley & Sons, Ltd.
Fivush, R., Haden, C., & Adam, S. (1995). Structure and coherence of preschoolers' personal narratives over time: Implications for childhood amnesia. *Journal of Experimental Child Psychology*, *60*, 32–56.
Fivush, R., Peterson, C., & Schwarzmueller, A. (2002). Questions and answers: The credibility of child witnesses in the context of specific questioning techniques. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview*. Mahwah, NJ: Erlbaum.
Flavell, J. H. (2000). Development of children's knowledge about the mental world. *International Journal of Behavioural Development*, *24*, 15–23.
Flavell, J. H., Botkin, P. T., Fry, C. K., Wright, J. C., & Jarvis, P. T. (1968). *The development of role-taking and communication skills in children*. New York: John Wiley & Sons, Inc.
Fleming, J. M., & Darhey, J. H. (1991). Mixed messages: The multiple audience problem and strategic communication. *Social Cognition*, *9*, 25–46.
Fleming, J. M., Darley, J. H., Hilton, B. A., & Kojetin, B. A. (1990). Multiple audience problem: A strategic communication perspective on social perception. *Journal of Personality and Social Psychology*, *58*, 593–609.

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development*, *56*, 1145–1155.

Ford, E. B. (2006). Lie detection: Historical, neuropsychiatric and legal dimensions. *International Journal of Law and Psychiatry*, *29*, 159–177.

Ford, C. V. (1995). *Lies! Lies!! Lies!!! The psychology of deceit*. Washington, DC: American Psychiatric Press.

Forrest, J. A., & Feldman, R. S. (2000). Detecting deception and judge's involvement: Lower task involvement leads to better lie detection. *Personality and Social Psychology Bulletin*, *26*, 118–125.

Forrest, J. A., Feldman, R. S., & Tyler, J. M. (2004). When accurate beliefs lead to better lie detection. *Journal of Applied Social Psychology*, *34*, 764–780.

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., & Ekman, P. (2004). Appearing truthful generalizes across different deception situations. *Journal of Personality and Social Psychology*, *86*, 486–495.

Frank, M. G., & Ekman, P., & Friesen, W. V. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of Personality and Social Psychology*, *64*, 83–93.

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

Frank, M. G., Feeley, T. H., Paolantonio, N., & Servoss, T. J. (2004). Individual and small group accuracy in judging truthful and deceptive communication. *Group Decision and Negotiation*, *13*, 45–59.

Frank, M. G., & Gilovich, T. (1988). The dark side of self- and social perception: Black uniforms and aggression in professional sports. *Journal of Personality and Social Psychology*, *54*, 74–85.

Freedman, J. L., Adam, E. K., Davey, S. A., & Koegl, C. J. (1996). The impact of a statement: More detail does not always help. *Legal and Criminological Psychology*, *1*, 117–130.

Freedman, N. (1972). The analysis of movement behavior during the clinical interview. In A. R. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 153–175). Elmsford, NY: Pergamon.

Freud, S. (1959). *Collected papers*. New York: Basic Books.

Friedman, H. S. (1978). The relative strength of verbal versus nonverbal cues. *Personality and Social Psychology Bulletin*, *4*, 147–150.

Friedman, H. S., & Miller-Herringer, T. (1991). Nonverbal display of emotion in public and private: Self-monitoring, personality and expressive cues. *Journal of Personality and Social Psychology*, *61*, 766–775.

Friedman, H. S., Riggio, R. E., & Casella, D. F. (1988). Nonverbal skill, personal charisma, and initial attraction. *Personality and Social Psychology Bulletin*, *14*, 203–211.

Friesen, W. V., Ekman, P., & Wallbott, H. (1979). Measuring hand movements. *Journal of Nonverbal Behavior*, *4*, 97–112.

Fugita, S. S., Wexley, K. N., & Hillery, J. M. (1974). Black-white differences in nonverbal behavior in an interview setting. *Journal of Applied Social Psychology*, *4*, 343–351.

Fukuda, K. (2001). Eye blinks: New indices for the detection of deception. *International Journal of Psychophysiology*, *40*, 239–245.

Furedy, J. J. (1991a). Alice in Wonderland terminological usage in, and communicational concerns about, that peculiarly flight of technological fancy: The CQT polygraph. *Integrative Physiological and Behavioral Science*, *26*, 241–247.

Furedy, J. J. (1991b). On the validity of the polygraph. *Integrative Physiological and Behavioral Science*, *26*, 211–213.

Furedy, J. J. (1993). The "control" question "test" (CQT) polygrapher's dilemma: Logico-ethical considerations for psychophysiological practioners and researchers. *International Journal of Psychophysiology*, *15*, 263–267.

Furedy, J. J. (1996a). Some elementary distinctions among, and comments concerning, the 'control' question 'test' (CQT) polygrapher's many problems: A reply to Honts, Kircher and Raskin. *International Journal of Psychophysiology*, *22*, 53–59.

Furedy, J. J. (1996b). The North American polygraph and psychophysiology: Disinterested, uninterested, and interested perspectives. *International Journal of Psychophysiology*, *21*, 97–105.

Gale, A. (1988). The polygraph test, more than scientific investigation. In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 1–9). London: Sage.

Galin, K. E., & Thorn, B. E. (1993). Unmasking pain: Detection of deception in facial expressions. *Journal of Social and Clinical Psychology*, *12*, 182–197.

Gallaher, P. E. (1992). Individual differences in nonverbal behavior: Dimensions of style. *Journal of Personality and Social Psychology*, *63*, 133–145.

Gallai, D. (1999). Polygraph evidence in federal courts: Should it be admissible? *American Criminal Law Review*, *36*, 87–116.

Gallup, G. (1982). Self-awareness and the emergence of mind in primates. *Americal Journal of Primatology*, *2*, 237–248.

Gamer, M., Rill, H. G., Vossel, G., & Gödert, H. W. (2006). Psychophysiological and vocal measures in the detection of guilty knowledge. *International Journal of Psychophysiology*, *60*, 76–87.

Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., & Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: An fMRI investigation. *Cerebral Cortex*, *13*, 830–836.

Garratt, G. A., Baxter, J. C., & Rozelle, R. M. (1981). Training university police in black-American nonverbal behavior. *The Journal of Social Psychology*, *113*, 217–229.

Garrido, E., & Masip, J. (2001). Previous exposure to the sender's behavior and accuracy at judging credibility. In R. Roesch, R. R. Corrado, & R. J. Dempster (Eds.), *Psychology in the courts: International advances in knowledge* (pp. 271–287). London: Routlegde.

Garrido, E., Masip, J., & Herrero, C. (2004). Police officers' credibility judgements: Accuracy and estimated ability. *International Journal of Psychology*, *39*, 254–275.

Geraerts, E. (2006). *Remembrance of things past. The cognitive psychology of remembering and forgetting trauma*. PhD thesis, Maastricht University.

Geraerts, E., Arnold, M. M., Lindsay, D. S., Merckelbach, H., Jelicic, M., & Hauer, B. (2006). Forgetting of prior remembering in people reporting recovered memories of child sexual abuse. *Psychological Science*, *17*, 1002–1008.

Geraerts, E., Schooler, J., Merckelbach, H., Jelicic, M., Hauer, B., & Ambadar, Z. (2007). The reality of recovered memories: Corroborating continuous and discontinuous memories of childhood sexual abuse. *Psychological Science, 18,* 564–568.

Ghetti, S., & Alexander, K. W. (2004). 'If it happened, I would remember it': Strategic use of event memorability in the rejection of false autobiographical events. *Child Development*, *75*, 542–561.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*, 107–119.

Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: Free Press.

Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, *75*, 332–346.

Gilstrap, L. L. (2004). A missing link in suggestibility research: What is known about the behavior of field interviewers in unstructured interviews with young children? *Journal of Experimental Psychology: Applied*, *10*, 13–24.

Ginton, A., Daie, N., Elaad, E., & Ben-Shakhar, G. (1982). A method for evaluating the use of the polygraph in a real life situation. *Journal of Applied Psychology*, *67*, 131–137.

Gödert, H. W., Gamer, M., Rill, H. G., & Vossel, G. (2005). Statement validity assessment: Inter-rater reliability of criteria-based content analysis in the mock-crime paradigm. *Legal and Crimiological Psychology*, *10*, 225–245.

Goffman, E. (1959). *The presentation of self in everyday life*. New York: Doubleday.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Doubleday.

Goodman, G. S., Batterman-Faunce, J. M., Schaaf, J. M., & Kenney, R. (2002). Nearly 4 years after an event: Children's eyewitness memory and adults' perceptions of children's accuracy. *Child Abuse & Neglect*, *26*, 849–884.

Goodman, G. S., & Melinder, A. (2007). Child witness research and forensic interviews of young children: A review. *Legal and Criminological Psychology, 12*, 1–20.

Goodman, G. S., Myers, J. E. B., Qin, J., Quas, J. A., Castelli, P., Redlich, A. D., & Rogers, L. (2006). Hearsay versus children's testimony: Effects of truthful and deceptive statements on jurors' decisions. *Law and Human Behavior*, *30*, 363–401.

Goodman, G. S., Rudy, L., Bottoms, B., & Aman, C. (1990). Children's concerns and memory: Issues of ecological validity in the study of children's eyewitness testimony. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 249–284). New York: Cambridge University Press.

Goodman, G. S., & Schwartz-Kenney (1992). Why knowing a child's age is not enough: Influences of cognitive, social, and emotional factors on children's testimony. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 15–32). Chichester, England: John Wiley & Sons, Ltd.

Gordon, A. K., & Miller, A. G. (2000). Perspective differences in the construal of lies: Is deception in the eye of the beholder? *Personality and Social Psychology Bulletin*, *26*, 46–55.

Gordon, B. N., Schroeder, C. S., & Abrams, J. M. (1990). Age and social class differences in children's knowledge of sexuality. *Journal of Clinical Child Psychology*, *19*, 33–43.

Gordon, N. J., & Fleisher, W. L. (2002). *Effective interviewing and interrogation techniques*. New York: Academic Press.

Gordon, R. A., Baxter, J. C., Rozelle, R. M., & Druckman, D. (1987). Expectations of honest, evasive and deceptive nonverbal behavior. *Journal of Social Psychology*, *127*, 231–233.

Gozna, L., & Babooram, N. (2004). Non-traditional interviews: Deception in a simulated customs baggage search. In A. Czerederecka, T. Jaskiewicz-Obydzinska, R. Roesch, & J. Wojcikiewicz (Eds,), *Forensic psychology and law* (pp. 153–161). Krakow, Poland: Institute of Forensic Research Publishers.

Gozna, L., & Forward, T. (2004). First impressions and stereotypical beliefs: An investigation of police officers' veracity judgements. In A. Czerederecka, T. Jaskiewicz-Obydzinska, R. Roesch, & J. Wojcikiewicz (Eds,), *Forensic psychology and law* (pp. 162-173). Krakow, Poland: Institute of forensic research publishers.

Gozna, L., Vrij, A., & Bull, R. (2001). The impact of individual differences on perceptions of lying in everyday life and in high stakes situation. *Personality and Individual Differences*, *31*, 1203–1216.

Granhag, P. A. (2006). Rethinking implicit lie detection. *Journal of Credibility Assessment and Witness Psychology*, *7*, 180–190.

Granhag, P. A., Andersson, L. O., Strömwall, L. A., & Hartwig, M. (2004). Imprisoned knowledge: Criminals' beliefs about deception. *Legal and Criminological Psychology*, *9*, 103–119.

Granhag, P. A., & Hartwig, M. (2007). *A new theoretical perspective on deception detection: On the psychology of instrumental mind reading*. Paper submitted for publication.

Granhag, P. A., & Strömwall, L. A. (1999). Repeated interrogations: Stretching the deception detection paradigm. *Expert Evidence: The International Journal of Behavioural Sciences in Legal Contexts*, *7*, 163–174.

Granhag, P. A., & Strömwall, L. A. (2000a). Effects of preconceptions on deception detection and new answers to why lie catchers often fail. *Psychology, Crime, & Law*, *6*, 197–218.

Granhag, P. A., & Strömwall, L. A. (2000b). "Let's go over this again...": Effects of repeated interrogations on deception detection performance. In A. Czerederecka, T. Jaskiewicz-Obydzinska, & J. Wojcikiewicz (Eds.), *Forensic psychology and law: Traditional questions and new ideas* (pp. 191–196). Krakow, Poland: Institute of Forensic Research Publishers.

Granhag, P. A., & Strömwall, L. A. (2001a). Deception detection: Examining the consistency heuristic. In C. M. Breur, M. M. Kommer, J. F. Nijboer, & J. M. Reijntjes (Eds.), *New trends in criminal investigation and evidence, volume 2* (pp. 309–321). Antwerpen, Belgium: Intresentia.

Granhag, P. A., & Strömwall, L. A. (2001b). Deception detection: Interrogators' and observers' decoding of consecutive statements. *Journal of Psychology*, *135*, 603–620.

Granhag, P. A., & Strömwall, L. A. (2001c). Deception detection based on repeated interrogations. *Legal and Criminological Psychology*, *6*, 85–101.

Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and nonverbal cues to deception. *Applied Cognitive Psychology*, *16*, 243–257.

Granhag, P. A., Strömwall, L. A., & Hartwig, M. (2005). Granting asylum or not? Migration board personnel's beliefs about deception. *Journal of Ethnic and Migration Studies*, *31*, 29–50.

Granhag, P. A., Strömwall, L. A. & Hartwig, M. (2007). The SUE technique: The way to interview to detect deception. *Forensic Update, 88,* January, 25–29.

Granhag, P. A., Strömwall, L. A., & Jonsson, A. C. (2003). Partners in crime: How liars in collusion betray themselves. *Journal of Applied Social Psychology*, *33*, 848–868.

Granhag, P. A., Strömwall, L. A., & Landström, S. (2006). Children recalling an event repeatedly: Effects on RM and CBCA scores. *Legal and Criminological Psychology*, *11*, 81–98.

Granhag, P. A., & Vrij, A. (2005). Deception detection. In N. Brewer & K. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 43–92). New York: Guilford Press.

Greene, J. O., O'Hair, H. D., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. *Human Communication Research*, *11*, 335–364.

Gregg, A. P. (2007). When vying reveals lying: The timed antagonistic response alethiometer. *Applied Cognitive Psychology, 21,* 621–648.

Greuel, L. (1992). Police officers' beliefs about cues associated with deception in rape cases. In F. Lösel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives* (pp. 234–239). Berlin: de Gruyter.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Griffin, D. W., & Bartholomew, K. (1994a). Models of the self and other: Fundamental dimensions underlying measurements of adult attachment. *Journal of Personality and Social Psychology*, *67*, 430–455.

Griffin, D. W., & Bartholomew, K. (1994b). The metaphysics of measurement: The case of adult attachment. *Advances in Personal Relationships*, *5*, 17–52.

Griffin, Z. M., & Oppenheimer, D. M. (2006). Speakers gaze at objects while preparing intentionally inaccurate labels for them. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 943–948.

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, *64*, 970–986.

Grubin, D., Madsen, L., Parsons, S., Sosnowski, D., & Warberg, B. (2004). A prospective study of the impact of polygraphy on high-risk behaviors in adult sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, *16*, 209–222.

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions*. Chichester, England: John Wiley & Sons, Ltd.

Gumpert, C. H., & Lindblad, F. (1999). Expert testimony on child sexual abuse: A qualitative study of the Swedish approach to statement analysis. *Expert Evidence*, *7*, 279–314.

Gumpert, C. H., Lindblad, F., & Grann, M. (2000). A systematic approach to quality assessment of expert testimony in cased of alleged child sexual abuse. *Psychology, Crime, & Law*, *8*, 59–75.

Gumpert, C. H., Lindblad, F., & Grann, M. (2002). The quality of written expert testimony in alleged child sexual abuse: An empirical study. *Psychology, Crime, & Law*, *8*, 77–92.

Hadjistavropoulos, H. D., & Craig, K. D. (1994). Acute and chronic low back pain: Cognitive, affective, and behavioral dimensions. *Journal of Consulting and Clinical Psychology*, *62*, 341–349.

Hadjistavropoulos, H. D., Craig, K. D., Hadjistavropoulos, T., & Poole, G. D. (1996). Subjective judgments of deception in pain expression: Accuracy and errors. *Pain*, *65*, 251–258.

Hala, S., & Russell, J. (2001). Executive control within strategic deception: A window on early cognitive development? *Journal of Experimental Child Psychology*, *80*, 112–141.

Hale, J. L., & Stiff, J. B. (1990). Nonverbal primacy in veracity judgments. *Communication Reports*, *3*, 75–83.

Hall, J. A. (1979). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*, 845–857.

Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and exppressive style*. Baltimore, MD: The Johns Hopkins University Press.

Hall, J. A. (2006). Women's and men's nonverbal communication. In V. Manusov & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 201–218). Thousand Oaks, CA: Sage.

Happel, M. D. (2005). Neuroscience and the detection of deception. *Review of Policy Research*, *22*, 667–685.

Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised* (2nd edition). Toronto, Ontario: Multi-Health Systems.

Hare, R. D., Forth, A. E., & Hart, S. D. (1989). The psychopath and prototype for pathological lying and deception. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 25–49). Dordrecht, the Netherlands: Kluwer Academic.

Harrigan, J. A., & O'Connell, D. M. (1996). Facial movements during anxiety states. *Personality and Individual Differences, 21,* 205–212.

Harrison, A. A., Hwalek, M., Raney, D. F., & Fritz, J. G. (1978). Cues to deception in an interview situation. *Social Psychology*, *41*, 156–161.

Hart, C. L., Hudson, L. P., Fillmore, D. G., & Griffith, J. D. (2006). Managerial beliefs about the behavioural cues of deception. *Individual Differences Research*, *4*, 176–184.

Hartwig, M. (2005). *Interrogation to detect deception and truth: Effects of strategic use of evidence*. PhD thesis, University of Gothenburg, Department of Psychology.

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

Hartwig, M., Granhag, P. A., Strömwall, L., & Andersson, L. O. (2004a). Suspicious minds: Criminals' ability to detect deception. *Psychology, Crime, & Law*, *10*, 83–95.

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

Hartwig, M., Granhag, P. A., Strömwall, L., & Vrij, A. (2004b). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly*, *7*, 429–456.

Hartwig, M., Granhag, P. A., Strömwall, L., & Vrij, A. (2005). Detecting deception via strategic closure of evidence. *Law and Human Behavior*, 29, 469–484.

Hartwig, M., Granhag, P. A., & Vrij, A. (2005). Police interrogation from a social psychology perspective. *Policing & Society*, *15*, 379–399.

Haselton, M. G., Buss, D. M., Oubaid, V., & Angleitner, A. (2005). Sex, lies, and strategic interference: The psychology of deception between sexes. *Personality and Social Psychology Bulletin*, *31*, 3–23.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, *52*, 511–524.

Heilveil, I., & Muehleman, J. T. (1981). Nonverbal clues to deception in a psychotherapy analogue. *Psychotherapy: Theory, Research and Practice*, *18*, 329–335.

Heinrich, C. U., & Borkenau, P. (1998). Deception and deception detection: The role of cross-modal inconsistency. *Journal of Personality*, *66*, 687–712.

Hemsley, G. D., & Doob, A. N. (1978). The effect of looking behavior on perceptions of a communicator's credibility. *Journal of Applied Social Psychology*, *8*, 136–144.

Henningsen, D. D., Cruz, M. G., & Morr, M. C. (2000). Pattern violations and perceptions of deception. *Communication Reports*, *13*, 1–9.

Hernandez-Fernaud, E., & Alonso-Quecuty, M. (1997). The cognitive interview and lie detection: A new magnifying glass for Sherlock Holmes? *Applied Cognitive Psychology*, *11*, 55–68.

Hershkowitz, I. (1999). The dynamics of interviews yielding plausible and implausible allegations of child sexual abuse. *Applied Developmental Science*, *3*, 28–33.

Hershkowitz, I. (2001a). A case study of child sexual false allegation. *Child Abuse & Neglect*, *25*, 1397–1411.

Hershkowitz, I. (2001b). Children's responses to open-ended utterances in investigative interviews. *Legal and Criminological Psychology*, *6*, 49–63.

Hershkowitz, I. (2002). The role of facilitative prompts in interviews of alleged sex and abuse victims. *Legal and Criminological Psychology*, *7*, 63–71.

Hershkowitz, I., Fisher, S., Lamb, M. E., & Horowitz, D. (2007). Improving credibility assessment in child sexual abuse allogations: The role of the NICHD investigative interview protocol. *Child Abuse & Neglect*, 31, 99–110.

Hershkowitz, I., Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1997). The relationships among interviewer utterance type, CBCA scores and the richness of children's responses. *Legal and Criminological Psychology*, *2*, 169–176.

Hess, U., & Kleck, R. E. (1990). Differentiating emotion elicited and deliberate emotional facial expressions. *European Journal of Social Psychology*, *20*, 369–385.

Hess, U., & Kleck, R. E. (1994). The cues decoders use in attempting to differentiate emotion-elicited and posed facial expressions. *European Journal of Social Psychology*, *24*, 367–381.

Hill, M. L., & Craig, K. D. (2002). Detecting deception in pain expressions: The structure of genuine and deceptive facial displays. *Pain*, *98*, 135–144.

Hira, S., Furumitsu, I., Nakayama, M. (2006). *The effects of retention intervals on detection of deception using P300*. Paper presented at the European Expert Meeting on Polygraph Testing. University of Maastricht, the Netherlands, 29–31 March.

Hirsch, A. R., & Wolf, C. J. (2001). Practical methods for detecting mendacity: A case study. *The Journal of the American Academy of Psychiatry and the Law*, *29*, 438–444.

Hocking, J. E., Bauchner, J., Kaminski, E. P., & Miller, G. R. (1979). Detecting deceptive communication from verbal, visual, and paralinguistic cues. *Human Communication Research*, *6*, 33–45.

Hocking, J. E., & Leathers, D. G. (1980). Nonverbal indicators of deception: A new theoretical perspective. *Communication Monographs*, *47*, 119–131.

Höfer, E., & Akehurst, L., & Metzger, G. (1996, August). *Reality monitoring: A chance for further development of CBCA?* Paper presented at the Annual meeting of the European Association on Psychology and Law in Siena, Italy.

Höfer, E., Köhnken, G., Hanewinkel, R., & Bruhn, C. (1993). *Diagnostik und attribution von glaubwürdigkeit*. Kiel: final report to the Deutsche Forschungsgemeinschaft, KO 882/4-2.

Hoffner, C., Cantor, J., & Thorson, E. (1989). Children's responses to conflicting auditory and visual features of a televised narrative. *Human Communication Research*, *16*, 256–278.

Holland, M. K., & Tarlow, G. (1972). Blinking and mental load. *Psychological Reports*, *31*, 119–127.

Holland, M. K., & Tarlow, G. (1975). Blinking and thinking. *Psychological Reports, 41,* 403–406.

Hollien, H. (1990). *The acoustics of crime: The new science of forensic phonetics*. New York: Plenum.

Holmberg, U., & Christianson, S. A. (2002). Murderers' and sexual offenders' experiences of police interviews and their inclination to admit or deny crimes. *Behavioral Sciences and the Law*, *20*, 31–45.

Honts, C. R. (1991). The emperor's new clothes: The application of the polygraph tests in the American workplace. *Forensic Reports*, *4*, 91–116.

Honts, C. R. (1994). Assessing children's credibility: Scientific and legal issues in 1994. *North Dakota Law Review*, *70*, 879–903.

Honts, C. R. (1995). The polygraph in 1995: Progress in science and the law. *North Dakota Law Review*, *17*, 987–1020.

Honts, C. R. (1996). Criterion development and validity of the CQT in field application. *Journal of General Psychology*, *123*, 309–324.

Honts, C. R. (2004). The psychophysiological detection of deception. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 103–123). Cambridge, England: Cambridge University Press.

Honts, C. R., & Alloway, W. R. (2007). Information does not affect the validity of a comparison question test. *Legal and Criminological Psychology, 12*, 311–320.

Honts, C. R., & Amato, S. L. (2002). Countermeasures. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 251–264). San Diego, CA: Academic Press.

Honts, C. R., Devitt, M. K., Winbush, M., & Kircher, J. C. (1996). Mental and physical countermeasures reduce the accuracy of the concealed knowledge test. *Psychophysiology*, *33*, 84–92.

Honts, C. R., Hodes, R. L., & Raskin, D. C. (1985). Effects of physical countermeasures on the physiological detection of deception. *Journal of Applied Psychology*, *70*, 177–187.

Honts, C. R., Kircher, J. C., & Raskin, D. C. (1996). Polygrapher's dilemma or psychologist's: A reply to Furedy's logico-ethical considerations for psychophysiological practitioners and researchers. *International Journal of Psychophysiology*, *20*, 199–207.

Honts, C. R., Kircher, J. C., & Raskin, D. C. (2002). The scientific status of research on polygraph techniques: The case for polygraph tests. In D. L. Faigman, D. Kaye, M. J. Saks, & J. Sanders (Eds.), *Modern scientific evidence: The law and science of expert testimony (Volume 2)* (pp. 446–483). St Paul, MN: West Law.

Honts, C. R., & Perry, M. V. (1992). Polygraph admissibility: Changes and challenges. *Law and Human Behavior*, *16*, 357–379.

Honts, C. R., Raskin, D. C., & Kircher, J. C. (1987). Effects of physical countermeasures and their electromyographic detection during polygraph tests for deception. *Journal of Psychophysiology*, *1*, 241–247.

Honts, C. R., Raskin, D. C., & Kircher, J. C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology*, *79*, 252–259.

Honts, C. R., Raskin, D. C., Kircher, J. C., & Hodes, R. L. (1988). Effects of spontaneous countermeasures on the physiological detection of deception. *Journal of Police Science and Administration*, *16*, 91–94.

Horowitz, S. W. (1991). Empirical support for statement validity assessment. *Behavioral Assessment*, *13*, 293–313.
Horowitz, S. W. (1998). Reliability of criteria-based content analysis of child witness statements: Response to Tully. *Legal and Criminological Psychology*, *3*, 189–193.
Horowitz, S. W., Lamb, M. E., Esplin, P. W., Boychuk, T. D., Krispin, O., & Reiter-Lavery, L. (1997). Reliabilty of criteria-based content analysis of child witness statements. *Legal and Criminological Psychology*, *2*, 11–21.
Horowitz, S. W., Lamb, M. E., Esplin, P. W., Boychuk, T. D., Reiter-Lavery, L., & Krispin, O. (1996). Establishing ground truth in studies of child sexual abuse. *Expert Evidence*, *4*, 42–52.
Horowitz, S. W., Kircher, J. C., Honts, C. R., & Raskin, D. C. (1997). The role of comparison questions in physiological detection of deception. *Psychophysiology*, *34*, 108–115.
Horvath, F. (1973). Verbal and nonverbal cues to truth and deception during polygraph examinations. *Journal of Police Science and Administration*, *1*, 138–152.
Horvath, F. (1978). An experimental comparison of the psychological stress evaluator and the galvanic skin response in detection of deception. *Journal of Applied Psychology*, *63*, 338–344.
Horvath, F. (1979). Effect of different motivational instructions on detection of deception with the psychological stress evaluator and the galvanic skin response. *Journal of Applied Psychology*, *64*, 323–330.
Horvath, F., Blair, J. P., & Buckley, J. P. (in press). The Behavioral Analysis Interview: Clarifying the practice, theory and understanding of its use and effectiveness. *International Journal of Police Science and Management*.
Horvath, F., Jayne, B., & Buckley, J. (1994). Differentiation of truthful and deceptive criminal suspects in behavioral analysis interviews. *Journal of Forensic Sciences*, *39*, 793–807.
Horvath, F., & 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.
Hunter, J. E., Gerbing, D. W., & Boster, F. J. (1982). Machiavellian beliefs and personality: Construct invalidity of the Machiavellianism dimension. *Journal of Personality and Social Psychology*, *43*, 1293–1305.
Hurd, K., & Noller, P. (1988). Decoding deception: A look at the process. *Journal of Nonverbal Behavior*, *12*, 217–233.
Iacono, W. G. (2000). The detection of deception. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology, 2nd edition* (pp. 772–793). Cambridge, England: Cambridge University Press.
Iacono, W. G., & Lykken, D. T. (1997). The validity of the lie detector: Two surveys of scientific opinion. *Journal of Applied Psychology*, *82*, 426–433.
Iacono, W. G., & Lykken, D. T. (1999) Update: The scientific status of research on the polygraph techniques: The case against polygraph tests. In D. L. Faigman, D. H. Kaye, M. J. Saks, & J. Sanders (Eds.), *Modern scientific evidence: The law and science of expert testimony* (Pocket Part, Volume 1) (pp. 174–184). St Paul, MN: West Law.
Ickes, W. (1984). Compositions in black and white: Determinants of interaction in interracial dyads. *Journal of Personality and Social Psychology*, *47*, 330–341.

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

Jackson, J. L., & Granhag, P. A. (1997). The truth or fantasy: The ability of barristers and laypersons to detect deception in children's testimony. In J. F. Nijboer & J. M. Reijntjes (Eds.), *Proceedings of the first world conference on new trends in criminal investigation and evidence* (pp. 213–220). Lelystad, The Netherlands: Koninklijke Vermande.

Jang, S. A., Smith, S. W., & Levine, T. R. (2002). To stay or to leave? The role of attachment styles in communication patterns and potential termination of romantic relationships following discovery of deception. *Communication Monographs*, *69*, 236–252.

Joffe, R., & Yuille, J. C. (1992, May). *Criteria-Based Content Analysis: An experimental investigation*. Paper presented at the NATO Advanced Study Institute on the child witness in context: Cognitive, social and legal perspectives, Lucca, Italy.

Johnson, A. K., Barnacz, A., Constantino, P., Triano, J., Shackelford, T. K., & Keenan, J. P. (2004). Female deception detection as a function of commitment and self-awareness. *Personality and Individual Differences*, *37*, 1417–1424.

Johnson, A. K., Barnacz, A., Yokkaichi, T., Rubio, J., Racioppi, C., Schackeford, T. K., Fisher, M. L., & Keenan, J. P. (2005). Me, myself, and lie: The role of self-awareness in deception. *Personality and Individual Differences*, *38*, 1847–1853.

Johnson, M. K. (1988). Reality Monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General*, *117*, 390–394.

Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues*, *40*, 33–50.

Johnson, M. K., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, *117*, 371–376.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–29.

Johnson, M. K., & Raye, C. L. (1981). Reality Monitoring. *Psychological Review*, *88*, 67–85.

Johnson, M. K., & Raye, C. L. (1998). False memories and confabulation. *Trends in Cognitive Sciences*, *2*, 137–146.

Johnson, M. M., & Rosenfeld, P. J. (1992). Oddball-evoked p300-based method of deception detection in the laboratory: II. Utilization of non-selective activation of relevant knowledge. *International Journal of Psychophysiology*, *12*, 289–306.

Johnson, R., Barnhardt, J., & Zhu, J. (2003). The deceptive response: Effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biological Psychology*, *64*, 217–253.

Johnson, R., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia*, *42*, 878–901.

Johnson, R., Barnhardt, J., & Zhu, J. (2005). Differential effects of practice on the executive processes used for truthful and deceptive responses: An event-related brain potential study. *Cognitive Brain Research*, *24*, 386–404.

Johnson, R. R. (2006). Confounding influences on police detection of suspiciousness. *Journal of Criminal Justice*, *34*, 435–442.

Jones, D. P. H., & Krugman, R. (1986). Can a three-year-old child bear witness to her sexual assault and attempted muder? *Child Abuse and Neglect*, *10*, 253–258.

Jones, D. P. H., & McGraw, J. M. (1987). Reliable and fictitious accounts of sexual abuse to children. *Journal of Interpersonal Violence*, *2*, 27–45.

Jones, D. P. H., & McQuinston, M. (1989). *Interviewing the sexually abused child*. London: Gaskell.

Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, *76*, 349–364.

Kalbfleisch, P. J. (1992). Deceit, distrust and the social milieu: Application of deception research in a troubled world. *Journal of Applied Communication Research*, 308–334.

Kalbfleisch, P. J. (1994). The language of detecting deceit. *Journal of Language and Social Psychology*, *13*, 469–496.

Kalbfleisch, P. J. (2001). Deceptive message intent and relational quality. *Journal of Language and Social Psychology*, *20*, 214–230.

Kalma, A., Witte, M., & Zaalberg, R. (1996). Authenticity: operationalization, manipulation, and behavioural components: An explaration. *Medium Psychologie*, *8*, 49–65.

Kaplar, M. E., & Gordon, A. K. (2004). The enigma of altruistic lying: Perspective differences in what motivates and justifies lie telling within romantic relationships. *Personal Relationships*, *11*, 489–507.

Kashy, D. A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology*, *70*, 1037–1051.

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

Kassin, S. M. (2004). True or false: "I'd know a false confession if I saw one". In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 172–194). Cambridge, England: Cambridge University Press.

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, *60*, 215–228.

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., Goldstein, C. J., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior*, *27*, 187–203.

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confessions: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*, 33–67.

Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior*, *29*, 211–227.

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

Kaufmann, G., Drevland, G. C., Wessel, E., Overskeid, G., & Magnussen, S. (2003). The importance of being earnest: Displayed emotions and witness credibility. *Applied Cognitive Psychology, 17,* 21–34.

Kebbel, M. R., & Milne, R. (1998). Police officers' perception of eyewitness factors in forensic investigations. *Journal of Social Psychology*, *138*, 323–330.

Keenan, J. P., Gallup, G., & Falk, D. (2003). *The face in the mirror*. New York: HarperCollins.

Keenan, J. P., Gallup, G. G., Goulet, N., & Kulkarni, M. (1997). Attriutions of deception in human mating strategies. *Journal of Social Behaviour and Personality*, *12*, 45–52.

Kemeny, M. E, (2003). The psychobiology of stress. *Current Directions in Psychological Science*, *4*, 124–129.

Kendall, S., & Tannen, D. (2001). Discourse and gender. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 548–567). Malden, MA: Blackwell.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, *27*, 175–200.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge, England: Cambridge University Press.

Kennedy, J., & Coe, W. C. (1994). Nonverbal signs of deception during posthypnotic amnesia. *The International Journal of Clinical and Experimental Hypnosis*, *17*, 13–19.

Kieras, J. E., Tobin, R. M., Braziano, W. G., & Rothbart, M. L. (2005). You can't always get what you want: Effortful control and children's responses to undesirable gifts. *Psychological Science*, *16*, 391–396.

Kim, H. S., & Baron, R. S. (1988). Exercise and the illusory correlation: Does arousal heighten stereotypical processing? *Journal of Experimental Social Psychology*, *24*, 366–380.

King, B. (2006). *The lying ape: An honest guide to a world of deception*. Cambridge, England: Icon books.

Kircher, J. C., Horowitz, S. W., & Raskin, D. C. (1988). Meta-analysis of mock crime studies of the control question polygraph technique. *Law and Human Behavior*, *12*, 79–90.

Kircher, J. C., & Raskin, D. C. (1988). Human versus computerized evaluations of polygraph data in a laboratory setting. *Journal of Applied Psychology*, *73*, 291–302.

Kircher, J. C., & Raskin, D. C. (2002). Computer methods for the psychophysiological detection of deception. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 287–326). San Diego, CA: Academic Press.

Klaver, J. R., Lee, Z., & Hart, S. D. (2007). Psychopathy and nonverbal indicators of deception in offenders. *Law and Human Behavior*, *31*, 337–351.

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

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*, 78–100.

Kleinmuntz, B., & Szucko, J. J. (1982). On the fallibility of lie detection. *Law and Society Review*, *17*, 85–104.

Kleinmuntz, B., & Szucko, J. J. (1984). Lie detection in ancient and modern times: A call for contemporary scientific study. *American Psychologist*, *39*, 766–776.

Kline, P. (1993). *The handbook of psychological testing*. New York: Routledge.

Knapp, M. L., Hart, R. P., & Dennis, H. S. (1974). An exploration of deception as a communication construct. *Human Communication Research*, *1*, 15–29.

Knight, J. (2004). The truth about lying. *Nature*, *428*, 692–694.

Knox, D., Zusman, M. E., McGinty, K., & Gescheidler (2001). Deception of parents during adolescence. *Adolescence*, *36*, 611–614.

Köhnken, G. (1985). Speech and deception of eyewitnesses: An information processing approach. In F. L. Denmark (Ed.), *The psychology of women* (pp. 117–139), North Holland, the Netherlands: Elsevier Science Publishers.

Köhnken, G. (1987). Training police officers to detect deceptive eyewitness statements. Does it work? *Social Behaviour*, 2, 1–17.

Köhnken, G. (1989). Behavioral correlates of statement credibility: Theories, paradigms and results. In H. Wegener, F. Lösel, & J. Haisch (Eds.), *Criminal behavior and the justice system: Psychological perspectives* (pp. 271–289). New York: Springer-Verlag.

Köhnken, G. (1996). Social psychology and the law. In G. R. Semin, & K. Fiedler (Eds.), *Applied social psychology* (pp. 257–282). London: Sage.

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. (2002). A German perspective on children's testimony. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 233–244). Chichester, England: John Wiley & Sons, Ltd.

Köhnken, G. (2004). Statement Validity Analysis and the 'detection of the truth'. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 41–63). Cambridge, England: Cambridge University Press.

Köhnken, G., Milne, R., Memon, A., & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime, and Law*, *5*, 3–28.

Köhnken, G., Schimossek, E., Aschermann, E., & Höfer, E. (1995). The cognitive interview and the assessment of the credibility of adult's statements. *Journal of Applied Psychology*, *80*, 671–684.

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, England: British Psychological Society.

Köhnken, G., & Wegener, H. (1982). Zur Glaubwürdigkeit von Zeugenaussagen: Experimentelle Uberprüfung ausgewählter Glaubwürdigkeitskriterien (Credibility of witness statements: Experimental examination of selected reality criteria). *Zeitschrift für Experimentelle und Angewandte Psychologie*, *29*, 92–111.

Kokish, R., Levenson, J. S., & Blasingame, G. D. (2005). Post-conviction sex offender polygraph examination: Client-reported perceptions of utility and accuracy. *Sexual Abuse: A Journal of Research and Treatment*, *17*, 211–221.

Kowalski, R. M., Walker, S., Wilkinson, R., Queen, A., & Sharpe, B. (2003). Lying, cheating, complaining, and other aversive interpersonal behaviors: A narrative examination of the darker side of relationships. *Journal of Social and Personal Relationships*, *20*, 471–490.

Kozel, F. A., Padgett, T. M., & George, M. S. (2004a). A replication study of the neural correlates of deception. *Behavioral Neuroscience*, *118*, 852–856.

Kozel, F. A., Revell, L. J., Lorberbaum, J. P., Shastri, A., Elhai, J. D., Horner, M. D., Smtih, A., Nahas, Z., Bohning, D. E., & George, M. S. (2004b). A pilot study of functional magnetic resonance imaging brain correlates of deception in healthy young men. *Journal of Neuropsychiatry and Clinical Neuroscience*, *16*, 295–305.

Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, *58*, 605–613.

Krapohl, D. J. (2002). The polygraph in personnel screening. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 217–236). San Diego, CA: Academic Press.

Krauss, R. M. (1981). Impression formation, impression management, and nonverbal behaviors. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium* (Vol 1, pp. 323–341). Hillsdale, NJ: Erlbaum.

Kraut, R. E. (1978). Verbal and nonverbal cues in the perception of lying. *Journal of Personality and Social Psychology*, *36*, 380–391.

Kraut, R. E. (1980). Humans as lie detectors: Some second thoughts. *Journal of Communication*, *30*, 209–216.

Kraut, R. E., & Poe, D. (1980). On the line: The deception judgments of customs inspectors and laymen. *Journal of Personality and Social Psychology*, *36*, 380–391.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review*, *103*, 262–283.

Kugelmass, S., & Lieblich, I. (1966). Effects of realistic stress and procedural interference in experimental lie detection. *Journal of Applied Psychology*, *50*, 211–216.

Kuhlman, M. S. (1980). Nonverbal communications in interrogations. *FBI Law Enforcement Bulletin*, *49*, 6–9.

Kurasawa, T. (1988). Effects of contextual expectancies on deception-detection. *Japanese Psychological Research*, *30*, 114–121.

LaFrance, M., & Mayo, C. (1976). Racial differences in gaze behavior during conversations: Two systematic observational studies. *Journal of Personality and Social Psychology*, *33*, 547–552.

LaFrance, M., & Mayo, C. (1978). Cultural aspects of nonverbal communication. *International Journal of Intercultural Relations*, *2*, 71–89.

Lakhani, M., & Taylor, R. (2003). Beliefs about the cues to deception in high- and low-stake situations. *Psychology, Crime, & Law*, *9*, 357–368.

Lamb, M. E. (1998). Mea culpa but caveat emptor! Response to Tully. *Legal and Criminological Psychology*, *3*, 193–195.

Lamb, M. E., Esplin, P. W., & Sternberg, K. J. (1995). Making children into competent witnesses: Reactions to the Amicus Brief in re Michaels. *Psychology, Public Policy, and Law*, *1*, 438–449.

Lamb, M. E., Hershkowitz, I., Sternberg, K. J., Esplin, P. W., Hovav, M., Manor, M., & Yudilevitch, L. (1996). Effects of investigative utterance types on Israeli children's responses. *International Journal of Behavioral Development*, *19*, 627–637.

Lamb, M. E., Orbach, Y., Sternberg, K. J., Esplin, P. W., & Hershkowitz, I. (2002). The effects of forensic interview practices on the quality of information provided by alleged victims of child abuse. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 131–145). Chichester: John Wiley and Sons, Ltd.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1994). Factors influencing the reliability and validity of statements made by young victims of sexual maltreatment. *Journal of Applied Developmental Psychology*, *15*, 255–280.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (1998). Conducting investigative interviews of alleged sexual abuse victims. *Child Abuse and Neglect*, *22*, 813–823.

Lamb, M. E., Sternberg, K. J., & Esplin, P. W. (2000). Effects of age and delay on the amount of information provided by alleged sex abuse victims in investigative interviews. *Child Development*, *71*, 1586–1596.

Lamb, M. E., Sternberg, K. J., Esplin, P. W., Hershkowitz, I., & Orbach, Y. (1997a). Assessing the credibility of children's allegations of sexual abuse: A survey of recent research. *Learning and Individual Differences*, *9*, 175–194.

Lamb, M. E., Sternberg, K. J., Esplin, P. W., Hershkowitz, I., Orbach, Y., & Hovav, M. (1997b). Criterion-based content analysis: A field validation study. *Child Abuse and Neglect*, *21*, 255–264.

Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., & Esplin, P. W. (1999). Forensic interviews of children. In R. Bull & A. Memon (Eds.), *The psychology of interviewing: A handbook* (pp. 253–277). Chichester, England: John Wiley & Sons, Ltd.

Lamers-Winkelman, F. (1995). *Seksueel misbruik van jonge kinderen: Een onderzoek naar signalen en signaleren, en naar ondervragen en vertellen inzake seksueel misbruik*. Amsterdam, the Netherlands: VU Uitgeverij.

Lamers-Winkelman, F. (1999). Statement Validity Analysis: Its application to a sample of Dutch children who may have been sexually abused. *Journal of Aggression, Maltreatment & Trauma*, *2*, 59–81.

Lamers-Winkelman, F., & Buffing, F. (1996). Children's testimony in the Netherlands: A study of Statement Validity Analysis. In B. L. Bottoms & G. S. Goodman (1996), *International perspectives on child abuse and children's testimony* (pp. 45–62). Thousand Oaks, CA: Sage.

Landry, K., & Brigham, J. C. (1992). The effect of training in Criteria-Based Content Analysis on the ability to detect deception in adults. *Law and Human Behavior*, *16*, 663–675.

Landström, S., Granhag, P. A., & Hartwig, M. (2005). Witnesses appearing live versus on video: Effects on observers' perception, veracity assessments and memory. *Applied Cognitive Psychology*, *19*, 913–933.

Landström, S., Granhag, P. A., & Hartwig, M. (2007). Children's live and videotaped testimonies: How presentation mode affects observers' perception, assessment and memory. *Legal and Criminological Psychology*, *12*, 333–348.

Lane, J. D., & DePaulo, B. M. (1999). Completing Coyne's Cycle: Dysphorics' ability to detect deception. *Journal of Research in Personality*, *33*, 311–329.

Langleben, D. D. (in press). Detection of deception with fMRI: Are we there yet? *Legal and Criminological Psychology*.

Langleben, D. D., Dattilio, F. M., & Guthei, T. G. (2006). True lies: Delusions and lie-detection technology. *Journal of Psychiatry & Law*, *34*, 351–370.

Langleben, D. D., Loughead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., & Gur, R. C. (2005). Telling the truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, *26*, 262–272.

Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., O'Brien, C. P., & Childress, A. R. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *NeuroImage*, *15*, 727–732.

Larochette, A. C., Chambers, C. T., & Craig, K. D. (2006). Genuine, suppressed and faked facial expression of pain in childen. *Pain*, 126, 64–71.

Larson, J. A. (1932). *Lying and its detection: A study of deception and deception tests*. Chicago, IL: University of Chicago Press.
Lawson, H. M., & Leck, K. (2006). Dynamics of internet dating. *Social Science Computer Review*, *24*, 189–208.
Lawson, T. J. (2000). Are kind lies better than unkind truths? Effects of the perspective and closeness of relationship. *Representative Research in Social Psychology*, *24*, 11–19.
Leach, A. M., Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2004). "Intuitive" lie detection of children's deception by law enforcement officials and university students. *Law and Human Behavior*, *28*, 661–685.
Leal, S. (2005). *Central and peripheral physiology of attention and cognitive demand: Understanding how brain and body work together*. PhD thesis, University of Portsmouth, Department of Psychology.
Leal, S., Vrij, A., Fisher, R., & van Hooff (2006). *Cognitively induced arousal suppression when lying*. Manusript submitted for publication.
Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*, 34–47.
Lee, K., & Cameron, C. A. (2000). Extracting truthful information from lies: Emergence of the expression-representation distinction. *Merrill-Palmer Quarterly*, *46*, 1–20.
Lee, T. M. C., Liu, H. L., Tan, L. H., Chan, C. C. H., Mahankali, S., Feng, C. M., Hou, J., Fox, P. T., & Gao, J. H. (2002). Lie detection by functional magnetic resonance imaging. *Human Brain Mapping*, 15, 157–164.
Leekam, S. R. (1992). Believing and deceiving: Steps to becoming a good liar. In S. J. Ceci, M. DeSimone Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 47–62). Hillsdale, NJ: Erlbaum.
Leippe, M. R., Brigham, J. C., Cousins, C., & Romanczyk, A. (1987). The opinions and practices of criminal attorneys regarding child eyewitnesses: A survey. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), *Perspectives on children's testimony* (pp. 100–130). New York: Springer-Verlag.
Leippe, M. R., Manion, A. P. & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fact finders judge the accuracy of adults' and children's memory reports? *Journal of Personality and Social Psychology*, *63*, 181–197.
Leo, R. A. (1996a). Inside the interrogation room. *Journal of Criminal Law and Criminology*, *86*, 266–303.
Leo, R. A. (1996b). Miranda's revenge. *Law and Society Review*, *30*, 259–288.
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.
Lerner, M. J. (1980). *The belief in a just world*. New York: Plenum.
Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment*, *14*, 299–316.
Levine, T. R. (2001). Dichotomous and continuous views of deception: A re-examination of deception ratings in Information Manipulation Theory. *Communication Research Reports*, *18*, 230–240.
Levine, T. R., Anders, L. N., Banas, J., Baum, K. L., Endo, K., Hu, A. D. S., & Wong, N. C. H. (2000). Norms, expectations, and deception: A norm violation model of veracity judgements. *Communication Monographs*, *67*, 123–137.
Levine, T. R., Asada, K. J. K., & Lindsey, L. L. M. (2003). The relative impact of violation type and lie severity on judgments of messages deceitfulness. *Communication Research Reports*, *20*, 208–218.

Levine, T. R., Asada, K. J. K., & Park, H. S. (2006). The lying chicken and the gaze avoidant egg: Eye contact, deception and causal order. *Southern Journal of Communication, 4,* 401–411.

Levine, T. R., Feeley, T. H., McCornack, S. A., Hughes, M., & Harms, C. M. (2005). Testing the effects of nonverbal behavior training on accuracy in deception detection with the inclusion of a bogus training control group. *Western Journal of Communication*, *69*, 203–217.

Levine, T. R., Kim, R. K., & Park, H. S., & Hughes, M. (2006). Deception detection accuracy is a predictable linear function of message veracity base-rate: A formal test of Park and Levine's probability model. *Communication Monographs*, *73*, 243–260.

Levine, T. R., & McCornack, S. A. (1992). Linking love and lies: A formal test of the McCornack and Parks model of deception detection. *Journal of Social and Personal Relationships*, *9*, 143–154.

Levine, T. R., & McCornack, S. A. (1996a). A critical analysis of the behavioral adaptation explanation of the probing effect. *Human Communication Research*, *22*, 575–588.

Levine, T. R., & McCornack, S. A. (1996b). Can behavioral adaptation explain the probing effect? *Human Communication Research*, *22*, 604–613.

Levine, T. R., & McCornack, S. A. (2001). Behavioural adaptation, confidence, and heuristic-based explanations of the probing effect. *Human Communication Research*, *27*, 471–502.

Levine, T. R., McCornack, S. A., & Avery, P. B. (1992). Sex differences in emotional reactions to discovered deception. *Communication Quarterly*, *40*, 289–296.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communication Monographs*, *66*, 125–144.

Lewis, M. (1993). The development of deception. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 90–105). New York: Guilford Press.

Lewis, M., Stanger, C., & Sullivan, M. W. (1989). Deception in three-year-olds. *Developmental Psychology*, *25*, 439–443.

Lind, E. A., Erickson, B. E., Conley, J., & O'Barr, W. M. (1978). Social attributions and conversation style in trial testimony. *Journal of Personality and Social Psychology*, *36*, 1558–1567.

Lindsay, D. S. (2002). Children's source monitoring. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 83–98). Chichester, England: John Wiley and Sons, Ltd.

Lindsay, D. S., & Johnson, M. K. (1987). Reality monitoring and suggestibility: Children's ability to discriminate among memories from different sources. In S. J. Ceci, J. Toglia, & D. F. Ross (Eds), *Children's eyewitness memory* (pp. 91–121). New York: Springer-Verlag.

Lindsay, R. C. L. (1994). Expectations of eyewitness performance: Jurors' verdicts do not follow from their beliefs. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: Current trends and developments* (pp. 362–284). New York: Cambridge University Press.

Lippard, P. V. (1988). "Ask me no questions, I'll tell you no lies": Situational exigencies for interpersonal deception. *Western Journal of Speech Communication*, *52*, 91–103.

Littlepage, G. E., & Pineault, M. A. (1985). Detection of deception of planned and spontaneous communications. *Journal of Social Psychology*, *125*, 195–201.

Littmann, E., & Szewczyk, H. (1983). Zu einigen Kriterien und Ergebnissen forensisch-psychologischer Glaubwürdigkeitsbegutachtung von sexuell misbrauchten Kindern und Jugendlichen. *Forensia*, *4*, 55–72.

Loftus, E. F., & Palmer, J. C. (1974). Reconstructions of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, *13*, 585–589.

Lorber, M. F. (2004). Psychophysiology of aggression, psychopathy, and conduct problems: A meta-analysis. *Psychological Bulletin*, *130*, 531–552.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098–2109.

Lucas, R., & McKenzie, I. K. (1998). The detection of dissimulation: Lies, damned lies and SVA. *International Journal of Police Science & Management*, *1*, 347–359.

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, *43*, 385–388.

Lykken, D. T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology*, *44*, 258–262.

Lykken, D. T. (1988). The case against polygraph testing. In A. Gale (Ed.), *The polygraph test: Lies, truth, and science* (pp. 111–126). London: Sage.

Lykken, D. T. (1991). Why (some) Americans believe in the lie detector while others believe in the Guilty Knowledge Test. *Integrative Physiological and Behavioral Science*, *126*, 214–222.

Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector*. New York: Plenum Press.

Macdonald, J.M., & Michaud, D.L. (1992). *Criminal interrogation*. Denver; Colorado: Apache Press.

MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *Journal of Applied Psychology*, *86*, 674–683.

Maclin, O. H., & Maclin, M. K. (2004). The alias advantage and perceptions of guilt. *Journal of Psychology*, *138*, 339–349.

MacNeil, B. M., & Holden, R. R. (2006). Psychopathy and the detection of faking on self-report inventories of personality. *Personality and Individual Differences*, *41*, 641–651.

Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: Categorical person perception. *British Journal of Psychology*, *92*, 239–256.

Magnussen, S., Andersson, J., Cornoldi, C., De Beni, R., Endestad, T., Goodman, G. S., Helstrup, T., Koriat, A., Larsson, M., Melinder, A., Nilsson, L. G., Rönnberg, R., & Zimmer, H. (2006). What people believe about memory. *Memory*, *14*, 595–613.

Maier, N. R. F., & Thurber, J. A. (1968). Accuracy of judgments of deception when an interview is watched, heard, and read. *Personnel Psychology*, *21*, 23–30.

Mak, E. G., Y., & Lee, T. M. C. (2006). Detection of feigned memory impairments using a Chinese word task. *Psychological Reports*, *98*, 779–788.

Malcolm, S. R., & Keenan, J. P. (2003). My right I: Deception detection and hemispheric differences in self-awareness. *Social Behavior and Personality*, *31*, 767–772.

Malcolm, S. R., & Keenan, J. P. (2005). Hemispheric asymmetry and deception detection. *Laterality*, *10*, 103–110.

Malone, B. E., & DePaulo, B. M. (2001). Measuring sensitivity to deception. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 103–124). Mahwah, NJ: Erlbaum.

Mann, S. (2001). *Suspects, lies and videotape: An investigation into telling and detecting lies in police/suspect interviews*. PhD thesis, University of Portsmouth, Department of Psychology.

Mann, S. & Vrij, A. (2006). Police officers' judgements of veracity, tenseness, cognitive load and attempted behavioural control in real life police interviews. *Psychology, Crime, & Law*, *12*, 307–319.

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.

Mann, S., Vrij, A., & Bull, R. (2006). Looking through the eyes of an accurate lie detector. *Journal of Credibility Assessment and Witness Psychology*, *7*, 1–16.

Mann, S., Vrij, A., Fisher, R., & Robinson M. (2007, published online), See no lies, hear no lies: Differences in discrimination accuracy and response bias when watching or listening to police suspect interviews. *Applied Cognitive Psychology*.

Manstead, A. S. R., Wagner, H. L., & MacDonald, C. J. (1986). Deceptive and nondeceptive communications: Sending experience, modality, and individual abilities. *Journal of Nonverbal Behavior*, *10*, 147–167.

Manzanero, A. L., & Digest, M. (1996). Effects of preparation on internal and external memories. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, Law, and Criminal Justice: International developments in research and practice* (pp. 56–63). Berlin: de Gruyter.

Markham, R. (1991). Development of reality monitoring for performed and imagined actions. *Perceptual and Motor Skills*, *72*, 1347–1354.

Marshall, B. C., & Alison, L. J. (2006). Structural behavioural analysis as a basis for discriminating between genuine and simulated rape allegations. *Journal of Investigative Psychology and Offender Profiling*, *3*, 21–34.

Marston, W. A. (1917). Systolic blood pressure symptoms of deception. *Journal of Experimental Psychology*, *2*, 117–163.

Masip, J. (2006). The social psychology of the detection of deception. *Law and Human Behavior*, *30*, 403–407.

Masip, J., Alonso, H., Garrido, E., Anton, C. (2005). Generalized communicative suspicion (GCS) among police officers: Accounting for the investigator bias effect. *Journal of Applied Social Psychology*, *35*, 1046–1066.

Masip, J., Garrido, E., & Herrero, C. (2003a). Facial appearance and judgments of credibility: The effects of facial babyishness and age on statement credibility. *Genetic, Social, and General Psychology Monographs*, *129*, 269–311.

Masip, J., Garrido, E., & Herrero, C. (2003b). When did you conclude she was lying? The impact of the moment the decision about the sender's veracity is made and the sender's facial appearance on police officers' credibility judgements. *Journal of Credibility Assessment and Witness Psychology*, *4*, 1–36.

Masip, J., Garrido, E., & Herrero, C. (2004). Facial appearance and impressions of credibility: The effects of facial babyishness and age on person perception. *International Journal of Psychology*, *39*, 276–289.

Masip, J., Garrido, E., & Herrero, C. (2006). Observers' decision moment in deception detection experiments: Its impact on judgement, accuracy, and confidence. *International Journal of Psychology*, *41*, 304–319.

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

Matarazzo, J. D., Wiens, A. N., Jackson, R. H., & Manaugh, T. S. (1970). Interviewee speech behavior under conditions of endogenously-present and exogenously-induced motivational states. *Journal of Clinical Psychology*, *26*, 17–24.

Matsumoto, D. (2006). Culture and nonverbal behavior. In V. Manusov & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 219–235). Thousand Oaks, CA: Sage.

Maxwell, G. M., Cook, M.W., & Burr, R. (1985). The encoding and decoding of liking from behavioral cues in both auditory and visual channels. *Journal of Nonverbal Behavior*, *9*, 239–264.

Mazur, M. A., & Ebesu Hubbard, A. S. (2004). "Is there something I should know?" Topic avoidant responses in parent-adolescent communication. *Communication Reports*, *17*, 27–37.

McClintock, C. C., & Hunt, R. G. (1975). Nonverbal indicators of affect and deception in an interview setting. *Journal of Applied Social Psychology*, *5*, 54–67.

McCornack, S. A. (1992). Information manipulation theory. *Communication Monographs*, *59*, 1–16.

McCornack, S. A. (1997). The generation of deceptive messages: Laying the groundwork for a viable theory of interpersonal deception. In J. O. Greene (Ed.), *Message production: Advances in communication theory* (pp. 91–126). Mahway, NJ: Erlbaum.

McCornack, S. A., & Levine, T. R. (1990a). When lies are uncovered: Emotional and relational outcomes of discovered deception. *Communication Monographs*, *57*, 119–138.

McCornack, S. A., & Levine, T. R. (1990b). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communication Monographs*, *57*, 219–230.

McCornack, S. A., & Parks, M. R. (1986). Detection deception and relational development: The other side of trust. In M. L. McLaughlin (Ed.), *Communication Yearbook 9* (pp. 377–389). Beverly Hills, CA: Sage.

McCornack, S. A., & Parks, M. R. (1990). What women know that men don't: Sex differences in determining truth behind deceptive messages. *Journal of Social and Personal Relationships*, *7*, 107–118.

McCroskey, J. C., & Mehrley, S. (1969). The effects of disorganization and nonfluency on attitude change and source credibility. *Speech Monographs*, *36*, 13–21.

McKinnon, M. E. (1982). A guide to nonverbal deception indicators. *Law and Order*, 53–58.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*, 350–371.

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

Mehrabian, A. (1971). Nonverbal betraying of feeling. *Journal of Experimental Research in Personality*, *5*, 64–73.

Mehrabian, A. (1972). *Nonverbal communication*. Chicago, IL: Aldine-Atherton.

Meijer, E. (2004). Leugendetectie en wetenschap: Vriend of vijand? *De Psycholoog*, *39*, 174–179.

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.

Meissner, C. A., & Kassin, S. M. (2004). "You're guilty, so just confess!" Cognitive and behavioural confirmation biases in the interrogation room. In D. Lassiter (Ed.), *Interrogations, confessions, and entrapment* (pp. 85–106). New York: Kluwer Academic/Plenum.

Mei-tai Fan, R., Wagner, H. L., & Manstead, A. S. R. (1995). Anchoring, familiarity, and confidence in the detection of deception. *Basic and Applied Social Psychology*, *17*, 83–96.

Melinder, A., Goodman, G. S., Eilertsen, D. E., & Magnussen, S. (2004). Beliefs about child witnesses: A survey of professionals. *Psychology, Crime, & Law*, *10*, 347–366.

Memon, A., & Bull, R. (1999). *The handbook of the psychology of interviewing*. Chichester, England: John Wiley & Sons, Ltd.

Memon, A., Vrij, A., & Bull, R. (2003*). Psychology and Law: Truthfulness, accuracy, and credibility, second edition*. Chichester, England: John Wiley & Sons, Ltd.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships*, *6*, 159–179.

Mikulincer, M., & Shaver, P. R., (2005). Attachment theory and emotions in close relationships: Exploring the attachment-related dynamics of emotional reactions to relational events. *Personal Relationships, 12,* 149–168.

Millar, M. G., & Millar, K. U. (1995). Detection of deception in familiar and unfamiliar persons: The effects of information restriction. *Journal of Nonverbal Behavior*, *19*, 69–84.

Millar, M. G., & Millar, K. U. (1997a). Effects of situational variables on judgments about deception and detection accuracy. *Basic and Applied Social Psychology*, *19*, 401–410.

Millar, M. G., & Millar, K. U. (1997b). The effects of cognitive capacity and suspicion on truth bias. *Communication Research*, *24*, 556–570.

Millar, M. G., & Millar, K. U. (1998). The effects of suspicion on the recall of cues used to make veracity judgments. *Communication Reports*, *11*, 57–64.

Miller, G. R., Bauchner, J. E., Hocking, J. E., Fontes, N. E., Kaminski, E. P., & Brandt, D. R. (1981). ". . . and Nothing but the truth". In B. D. Sales (Ed.), *Perspectives in Law and Psychology, volume II: The trial process* (pp. 145–179). New York: Plenum.

Miller, G. R., deTurck, M. A., & Kalbfleisch, P. J. (1983). Self-Monitoring, rehearsal, and deceptive communication. *Human Communication Research*, *10*, 97–117.

Miller, G. R., Mongeau, P. A., & Sleight, C. (1986). Fudging with friends and lying to lovers: Deceptive communication in personal relationships. *Journal of Social and Personal Relationships*, *3*, 495–512.

Miller, G. R., & Stiff, J. B. (1993). *Deceptive Communication*. Newbury Park, CA: Sage.

Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice*. Chichester, England: John Wiley & Sons, Ltd.

Milne, R., & Bull, R. (2003). Does the cognitive interview help children to resist the effects of suggestive interviewing? *Legal and Criminological Psychology*, *8*, 21–38.

Milne, R., & Bull, R. (2006). Interviewing victims of crime, including children and people with intellectual disabilities. In M. R. Kebbell & G, M. Davies

(Eds.), *Practical psychology for forensic investigations and prosecutions* (pp. 103–119). Chichester, England: John Wiley & Sons, Ltd.

Mitchell, D. C. (2002). *The pre-test interview: A preliminary framework*. In M. Kleiner (Ed.), Handbook of polygraph testing (pp. 183–216). San Diego, CA: Academic Press.

Mitchell, R. W. (1986). A framework for discussing deception. In R. W. Mitchell & N. S. Mogdil (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 3–40). Albany: State University of New York Press.

Mohamed, F. B., Faro, S. H., Gordon, N. J., Platek, S. M., Ahmad, H., & Williams, J. M. (2006). Brain mapping of deception and truth telling about an ecologically valid situation: Functional MR imaging and polygraph investigation - initial experience. *Radiology*, *238*, 679–688.

Monahan, J. (1995). Information processing differences of conversational participants and observers: The effects of self-presentation concerns and cognitive load. *Communication Monographs*, *62*, 265–281.

Morency, N. L., & Krauss, R. M. (1982). Children's nonverbal encoding and decoding of affect. In R. S. Feldman (Ed.), *Development of nonverbal behaviour in children* (pp. 181–199). New York: Springer-Verlag.

Moston, S. (1987). The suggestibility of children in interview studies. *Child Language*, *7*, 67–78.

Moston, S. (1992). Truth or lies. *Policing*, *8*, 26–39.

Moston, S. J., & Engelberg, T. (1993). Police questioning techniques in tape recorded interviews with criminal suspects. *Policing and Society*, *6*, 61–75.

Moston, S. J., Stephenson, G. M., & Williamson, T. M. (1992). The effects of case characteristics on suspect behaviour during police questioning. *British Journal of Criminology*, *32*, 23–39.

Moston, S. J., Stephenson, G. M., & Williamson, T. M. (1993). The incidence, antecedents and consequences of the use of the right to silence during police questioning. *Criminal Behaviour and Mental Health*, *3*, 30–47.

Motley, M. T. (1974). Accoustic correlates of lies. *Western Speech*, *38*, 81–87.

Mulder, M., & Vrij, A. (1996). Explaining conversation rules to children: An intervention study to facilitate children's accurate responses. *Child Abuse & Neglect*, *20*, 623–631.

Nakayama, M. (2002). Practical use of the concealed information test for criminal investigation in Japan. In M. Kleiner (Ed.). *Handbook of polygraph testing* (pp. 49–86). San Diego, CA: Academic Press.

National Research Council (2003). *The polygraph and lie detection*. Committee to Review the Scientific Evidence on the Polygraph. Washington, DC: The National Academic Press.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. N. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, *29*, 665–675.

Newton, P., Reddy, V., & Bull, R. (2000). Children's everyday deception and performance on false-belief tasks. *British Journal of Developmental Psychology*, *18*, 297—317.

Nigro, G. N., Buckley, M. A., Hill, D. E., & Nelson, J. (1989). When juries 'hear' children testify: the effects of eyewitness age and speech style on juror's perceptions of testimony. In S. J. Ceci, D. E. Ross, & M. P. Toglia (Eds.), *Perspectives on children's testimony* (pp. 57–70). New York: Springer-Verlag.

Nigro, G. N., & Snow, A. L. (1992). Sex, lies, and smiling faces: A brief report on gender differences in 3-year-olds' deceptions. In S. J. Ceci, M. DeSimone

Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 63–68). Hillsdale, NJ: Erlbaum.
Noller, P. (1985). Video primacy – a further look. *Journal of Nonverbal Behavior*, *9*, 28–47.
Nunez, J. M., Casey, B. J., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neutral substrates with response conflict and cognitive control. *NeuroImage*, *25*, 267–277.
O'Connell, D. C., Kowall, S., & Dill, E. J. (2004). Dialogicality in TV news interviews. *Journal of Pragmatics*, *36*, 185–205.
Ofshe, R. (1989). Coerced confessions: The logic of seemingly irrational action. *Cultic Studies Journal*, *6*, 1–15.
Ofshe, R. J., & Leo, R. A. (1997). The decision to confess falsely: Rational choice and irrational action. *Denver University Law Review*, *74*, 979–1112.
O'Hair, H. D., Cody, M., & McLaughlin, M. L. (1981). Prepared lies, spontaneous lies, Machiavellianism and nonverbal communication. *Human Communication Research*, *7*, 325–339.
Oldershaw, L., & Bagby, R. M. (1997). Children and deception. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 153–166).
Olsen, D. E., Harris, J. C., Capps, M. H., & Ansley, N. (1997). Computerized polygraph scoring system. *Journal of Forensic Sciences*, *42*, 61–70.
Orbach, Y., & Lamb, M. E. (1999). Assessing the accuracy of child's account of sexual abuse: A case study. *Child Abuse & Neglect*, *23*, 91–98.
Orcutt, H. K., Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., & Thomas, S. (2001). Detecting deception in children's testimony: Fact finders' abilities to reach the truth in open court and closed-circuit trials. *Law and Human Behavior*, *25*, 339–372.
Osgood, C. E. (1960). Some effects of motivation on style of encoding. In T. Sebok (Ed.), *Style in language* (pp. 293–306). Cambridge, MA: MIT Press.
O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The body-who-cried-wolf-effect. *Personality and Social Psychology Bulletin*, *29*, 1316–1327.
O'Sullivan, M. (2005). Emotional intelligence and deception detection: Why most people can't "read" others, but a few can. In R. E. Riggio & R. S. Feldman (Eds.), *Applications of nonverbal communication* (pp. 215–253). Mahwah, NJ: Erlbaum.
O'Sullivan, M. (2007). Unicorns or Tiger Woods: Are lie detection experts myths or rarities? A response to on lie detection "wizards" by Bond and Uysal. *Law and Human Behavior*, *31*, 117–123.
O'Sullivan, M., & Ekman, P. (2004). The wizards of deception detection. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 269–286). Cambridge, England: Cambridge University Press.
O'Sullivan, M., Ekman, P., & Friesen, W. V. (1988). The effect of comparisons on detecting deceit. *Journal of Nonverbal Behavior*, *12*, 203–216.
Ost, J., Vrij, A., Costall, A., & Bull, R. (2002). Crashing memories and reality monitoring: Distinguishing between perceptions, imaginations and 'false memories'. *Applied Cognitve Psychology*, *16*, 125–134.
O'Toole, D., Yuille, J. C., Patrick, C. J., & Iacono, W. G. (1994). Alcohol and the physiological detection of deception: Arousal and memory influences. *Psychophysiology*, *31*, 253–263.
Park, H. S., & Levine, T. R. (2001). A probability model of accuracy in deception detection experiments. *Communication Monographs*, *68*, 201–210.

Park, H. S., Levine, T. R., McCornack, S. A., Morrisson, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs*, *69*, 144–157.

Parker, A. D., & Brown, J. (2000). Detection of deception: Statement validity analysis as a means of determining truthfulness or falsity of rape allegations. *Legal and Criminological Psychology*, *5*, 237–259.

Parker, J. F. (1995). Age differences in source monitoring of performed and imagined actions on immediate and delayed tests. *Journal of Experimental Child Psychology*, *60*, 84–101.

Parliament, L., & Yarmey, A. D. (2002). Deception in eyewitness identification. *Criminal Justice and Behavior*, *29*, 734–746.

Patrick, C. J., & Iacono, W. G. (1989). Psychopathy, threat, and polygraph test accuracy. *Journal of Applied Psychology*, *74*, 347–355.

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.

Patterson, M. L. (1995). Invited article: A parallel process model of nonverbal communication. *Journal of Nonverbal Behavior*, *19*, 3–29.

Patterson, M. L. (2006). The evolution of theories of interactive behavior. In V. Manusov & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 21–39). Thousand Oaks, CA: Sage.

Pavlidis, J., Eberhardt, N. L., & Levine, J. A. (2002a). Seeing through the face of deception. *Nature*, *415*, 35.

Pavlidis, J., Eberhardt, N. L., & Levine, J. A. (2002b). Erratum: Seeing through the face of deception. *Nature*, *415*, 602.

Pavlov, I. P. (1927). *Condition reflex*. Oxford, England: Clarendon Press.

Pearson, H. (2006). Lure of lie detectors spooks ethicists. *Nature*, *441*, 918–919.

Pelham, B. W., & Neter, E. (1995). The effect of motivation on judgment depends on the difficulty of the judgment. *Journal of Personality and Social Psychology*, *68*, 581–594.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC 2001 manual*. Mahwah, NJ: Erlbaum.

Pennebaker, J. W., & Graybeal, A. (2001). Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science*, *10*, 90–93.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Languge use as an individual difference. *Journal of Personality and Social Psychology*, *77*, 1296–1312.

Peskin, J. (1992). Ruse and representations: On children's ability to conceal information. *Developmental Psychology*, *28*, 84–89.

Peters, D. P. (1991). The influence of stress and arousal on the child witness. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 60–76). Washington, DC: American Psychological Association.

Peterson, C. C. (1995). The role of perceived intention to deceive in children's and adults' concepts of lying. *British Journal of Developmental Psychology*, *13*, 237–260.

Pezdek, K., & Hinz, T. (2002). The construction of false events in memory. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 99–116). Chichester, England: John Wiley & Sons, Ltd.

Pezdek, K., & Hodge, D. (1999). Planting false childhood memories in children: The role of event plausibility. *Child Development*, *70*, 887–895.

Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., Bidrose, S., Pipe, M. E., Rogers, M., & Brodie, L. (2004). Detecting deception in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology*, *89*, 119–126.

Pezdek, K., & Roe, C. (1997). The suggestibility of children's memory for being touched: Planting, erasing, and changing memories. *Law and Human Behavior*, *21*, 95–106.

Pezdek, K., & Taylor, J. (2000). Discriminating between accounts of true and false events. In D. F. Bjorklund (Ed.), *Research and theory in false memory creation in children and adults* (pp. 69–91). Mahwah, NJ: Erlbaum.

Phan, K. L., Magalhaes, A., Ziemlewicz, T. J., Fitzgerald, D. A., Green, C., & Smith, W. (2005). Neural correlates of telling lies: A functional magnetic resonance imaging study at 4 Tesla. *Academic Radiology*, *12*, 164–172.

Pickel, K. L. (2004). When a lie becomes the truth: The effects of self-generated misinformation on eyewitness memory. *Memory*, *12*, 14–26.

Podlesny, J. A. (1993). Is the guilty knowledge polygraph technique applicable in criminal investigations? A review of FBI case records. *Crime Laboratory Digest*, *20*, 57–61.

Podlesny, J. A., & Raskin, D. C. (1977). Physiological measures and the detection of deception. *Psychological Bulletin*, *84*, 782–799.

Polace, D. C. (2004). Fabrication deflation? The mixed effects of lying on memory. *Applied Cognitive Psychology*, *18*, 455–465.

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.

Pontari, B. A., & Schlenker, B. R. (2006). Helping friends manage impressions: We like helpful liars but respect nonhelpful truth tellers. *Basic and Applied Social Psychology*, *28*, 177–183.

Poole, D. A., & White, L. T. (1991). Effects of question repetition and retention interval on the eyewitness testimony of children and adults. *Developmental Psychology*, *27*, 975–986.

Porter, S., Campbell, M. A., Stapleton, J., & Birt, A. R. (2002). The influence of judge, target, and stimulus characteristics on the accuracy of detecting deceit. *Canadian Journal of Behavioural Science*, *34*, 172–185.

Porter, S., Doucette, N. L., Woodworth, M., Earle, J., & MacNeil, B. (in press). "Halfe the world knows not how the other halfe lies": Investigation of cues to deception exhibited by criminal offenders and non-offenders. *Legal and Criminological Psychology*.

Porter, S., & Woodworth, M. (2007). "I'm sorry I did it...but he started it": A comparison of the official and self-reported homicide descriptions of psychopaths and non-psychopaths. *Law and Human Behavior, 31*, 91–107.

Porter, S., Woodworth, M., & Birt, A. R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and Human Behavior*, *24*, 643–658.

Porter, S., Woodworth, M., McCabe, S., & Peace, K. A. (2007). "Genius is 1% inspiration and 99% perspiration"...or is it? An investigation of the impact of motivation and feedback on deception detection. *Legal and Criminological Psychology, 12*, 297–310.

Porter, S., & Yuille, J. C. (1995). Credibility assessment of criminal suspects through statement analysis. *Psychology, Crime, & Law*, *1*, 319–331.

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior*, *20*, 443–459.

Porter, S., Yuille, J. C., & Birt, A. R. (2001). The discrimination of deceptive, mistaken, and truthful witness testimony. In R. Roesch, R. R. Corrado, & R. Dempster (Eds.), *Psychology in the courts: International advances in knowledge* (pp. 253–270). London: Routledge.

Porter, S., Yuille, J. C., & Lehman, D. R. (1999). The nature of real, implanted and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law and Human Behavior*, *23*, 517–537.

Pynoos, R. S., & Eth, S. (1984). The child as witness to homocide. *Journal of Social Issues*, *40*, 87–108.

Pynoos, R. S., & Nader, K. (1988). Children who witness the sexual assaults of their mothers. *Journal of the American Academy of Child and Adolescent Psychiatry*, *27*, 567–572.

Quas, J. A., Davis, E. L., Goodman, G. S., & Myers, J. E. B. (2007). Repeated questions, deception, and children's true and false reports of body touch. *Child Maltreatment*, *12*, 60–67.

Rabon, D. (1992). *Interviewing and interrogation*. Durham, North Carolina: Carolina Academic Press.

Raskin, D. C. (1979). Orienting and defensive reflexes in the detection of deception. In H. D. Kimmel, E. H. Van Olst, & J. F. Orlebeke (Eds.), *The orienting reflex in humans* (pp. 587–605). Hillsdale, NJ: Erlbaum.

Raskin, D. C. (1982). The scientific basis of polygraph techniques and their uses in the judicial process. In A. Trankell (Ed.), *Reconstructing the past* (pp. 317–371). Stockholm, Sweden: Norsted & Soners.

Raskin, D. C. (1986). The polygraph in 1986: Scientific, professional, and legal issues surrounding acceptance of polygraph evidence. *Utah Law Review*, *29*, 29–74.

Raskin, D. C. (1988). Does science support polygraph testing? In A. Gale (Ed.), *The polygraph test: Lies, truth and science* (pp. 96–110). London: Sage.

Raskin, D. C. (1989). Polygraph techniques for the detection of deception. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 247–296). New York: Springer Verlag.

Raskin, D. C., & Esplin, P. W. (1991a). Assessment of children's statements of sexual abuse. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 153–165). Washington, DC: American Psychological Association.

Raskin, D. C., & Esplin, P. W. (1991b). Statement Validity Assessment: Interview procedures and content analysis of children's statements of sexual abuse. *Behavioral Assessment*, *13*, 265–291.

Raskin, D. C., & Hare, R. D. (1978). Psychopathy and detection of deception in a prison population. *Psychophysiology*, *15*, 126–136.

Raskin, D. C., & Honts, C. R. (2002). The comparison question test. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 1–47). San Diego, CA: Academic Press.

Raskin, D. C., Honts, C. R., & Kircher, J. C. (1997). The scientific status of research on polygraph techniques: The case for polygraph tests. In D. L. Faigman, D. H. Kaye, M. J. Saks, & J. Sanders (Eds.), *Modern scientific evidence: The law and science of expert testimony* (pp. 565–582). St Paul, MN: West Law.

Raskin, D. C., Kircher, J. C., Horowitz, S. W., & Honts, C. R. (1989). Recent laboratory and field research on polygraph techniques. In J. C. Yuille (Ed.), *Credibility Assessment* (pp. 1–24). Dordrecht, the Netherlands: Kluwer Academic.

Raskin, D. C., & Steller, M. (1989). Assessing the credibility of allegations of child sexual abuse: Polygraph examinations and statement analysis. In H. Wegener, F. Losel, & J. Haisch (Eds.), *Criminal behavior and the justice system* (pp. 290–302). New York: Springer.

Raskin, D. C., & Yuille, J. C. (1989). Problems in evaluating interviews of children in sexual abuse cases. In S. J. Ceci, D. F. Ross, & M. P. Toglia (Eds.), *Perspectives on children's testimony* (pp. 184–207). New York: Springer.

Rassin, E. (1999). Criteria-Based Content Analysis: The less scientific road to truth. *Expert Evidence*, *7*, 265–278.

Rassin, E., & van der Sleen, J. (2005). Characteristics of true versus false allegations of sexual offences. *Psychological Reports*, *97*, 589–598.

Reddy, V. (2007). Getting back to the rough ground: Deception and "social living". *Philosophical Transactions of the Royal Society B-Biological Sciences*, *362*, 621–637.

Reid, J. E. (1947). A revised questioning technique in lie detection tests. *Journal of Criminal Law, Criminology, and Police Science*, *37*, 542–547.

Reid, J. E., & Arther, R. O. (1953). Behavior symptoms of lie-detector subjects. *Journal of Criminal Law, Criminology and Police Science*, *44*, 104–108.

Reid, J. E., & Inbau, F. E. (1977). *Truth and deception: The polygraph (lie detector) technique*. Baltimore: Williams & Wilkins.

Reis, H. T., Senchak, M., & Solomon, B. (1985). Sex differences in the intimacy of social interaction: Further examination of potential explanations. *Journal of Personality and Social Psychology*, *48*, 1204–1217.

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology*, *51*, 649–660.

Riggio, R. E. (1994). Epilogue: Where are we going, and how do we get there? *Journal of Language and Social Psychology*, *13*, 514–518.

Riggio, R. E., (2006). Nonverbal skills and abilities. In V. Manusov & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 79–95). Thousand Oaks, CA: Sage.

Riggio, R. E., & Friedman, H. S. (1983). Individual differences and cues to deception. *Journal of Personality and Social Psychology*, *45*, 899–915.

Riggio, R. E., Tucker, J., & Throckmorton, B. (1988). Social skills and deception ability. *Personality and Social Psychology Bulletin*, *13*, 568–577.

Riggio, R. E., Tucker, J., & Widaman, K. F. (1987). Verbal and nonverbal cues as mediators of deception ability. *Journal of Nonverbal Behavior*, *11*, 126–145.

Rime, B., & Schiaratura, L. (1991). Gesture and speech. In B. Rime & R. S. Feldman (Eds.), *Fundamentals of nonverbal behavior* (pp. 239–281). New York: Cambridge University Press.

Roberts, K. P., Lamb, M. E., Zale, J. L., & Randall, D. W. (1998). *Qualitative differences in children's accounts of confirmed and unconfirmed incidents of sexual abuse*. Paper presented at the biennial meeting of the American Psychology-Law Society, Redondo Beach, 5–7 March.

Robinson, W. P., (1994). Reactions to falsifications in public and and interpersonal contexts. *Journal of Language and Social Psychology*, *13*, 497–513.

Robinson, W. P., Shepherd, A., & Heywood, J. (1998). Truth, equivocation/concealment, and lies in job applications and doctor-patient communication. *Journal of Language and Social Psychology*, *17*, 149–164.

Rockwell, P., Buller, D. B., & Burgoon, J. K. (1997). Measurement of deceptive voices: Comparing acoustic and perceptual data. *Applied Psycholinguistics*, *18*, 471–484.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language*, *35*, 76–100.

Rosenfeld, J. P. (1995). Alternative views of Bashore and Rapp's (1993) alternative to traditional polygraphy: A critique. *Psychological Bulletin*, *117*, 159–166.

Rosenfeld, J. P. (2002). Event-related potential in the detection of deception, malingering, and false memories. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 265–286). San Diego, CA: Academic Press.

Rosenfeld, P. J., Angell, A., Johnson, M. M., & J. Qian (1991). An ERP-based, control-question lie detector analog: Algorithms for discriminating effects within individuals' average waveforms. *Psychophysiology*, *28*, 319–335.

Rosenfeld, P. J., Biroschak, J. R., & Furedy, J. J. (2006). P300-based detection of concealed autobiopgraphical versus incidentally acquired information in target and non-target paradigms. *International Journal of Psychophysiology*, *60*, 251–259.

Rosenfeld, J. P., Cantwell, B., Nasman, V. T., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience*, *24*, 157–161.

Rosenfeld, J. P., Ellwanger, J. W., Nolan, K., Wu, S., Bermann, R. G., & Sweet, J. (1999). P300 scalp amplitude distribution as an index of deception in a simulated cognitive deficit model. *International Journal of Psychophysiology*, *33*, 3–19.

Rosenfeld, J. P., Ellwanger, J., & Sweet, J. (1995). Detecting simulated amnesia with event-related brain potentials. *International Journal of Psychophysiology*, *19*, 1–11.

Rosenfeld, J. P., Rao, A., Soskins, M., & Miller, A. R. (2003). Scaled P300 scalp distribution correlates of verbal deception in an autobiographical oddball paradigm: Control for task demand. *Journal of Psychophysiology*, *17*, 14–22.

Rosenfeld, J. P., Reinhart, A. M., Bhatt, M., Ellwanger, J., Gora, K., Sekera, M., & Sweet, J. (1998). P300 correlates of simulated malingered amnesia in a matching-to-sample task: Topographic analyses of deception versus truthtelling responses. *International Journal of Psychophysiology*, *28*, 233–247.

Rosenfeld, J. P., Shue, E., & Singer, E. (2007). Single versus multiple blocks of P300-based concealed information tests for self-referring versus incidentally obtained information. *Biological Psychology*, *74*, 396–404.

Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, *41*, 205–219.

Rosenthal, R. (1976). *Experimenter effects in behavioural research*. New York: Irvington.

Rosenthal, R., & DePaulo, B. M. (1979). Sex differences in eavesdropping on nonverbal cues. *Journal of Personality and Social Psychology*, *37*, 273–285.

Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *The Behavioral and Brain Sciences*, *3*, 377–415.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental psychology* (Vol 10, pp. 174–221). New York: Academic Press.

Rotenberg, K. J., Simourd, L., & Moore, D. (1989). Children's use of a verbal-nonverbal consistency principle to infer truth from lying. *Child Development*, *60*, 309–322.

Rotenberg, K. J., & Sullivan, C. (2003). Children's use of gaze and limb movement cues to infer deception. *Journal of Genetic Psychology*, *164*, 175–187.

Rothwell, J., Bandar, Z., O'Shea, J., & McLean, D. (2006). Silent talker: A new computer-based system for the analysis of facial cues to deception. *Applied Cognitive Psycholgy*, *20*, 757–777.

Rowatt, W. C., Cunningham, M. R., & Druen, P. B. (1998). Deception to get a date. *Personality and Social Psychology Bulletin*, *24*, 1228–1242.

Rowatt, W. C., Cunningham, M. R., & Druen, P. B. (1999). Lying to get a date: The effect of facial physical attractiveness on the willingness to deceive prospective dating partners. *Journal of Social and Personal Relationships*, *16*, 209–223.

Rozelle, R. M., & Baxter, J. C. (1975). Impression formation and danger recognition in experienced police officers. *Journal of Social Psychology*, *96*, 53–63.

Rozelle, R. M., & Baxter, J. C. (1978). The interpretation of nonverbal behavior in a role-defined interaction sequence: The police-citizen encounter. *Environmental Psychology and Nonverbal Behavior*, *2*, 167–181.

Ruback, R. B. (1981). Perceived honesty in the parole interview. *Personality and Social Psychology Bulletin*, *7*, 677–681.

Ruback, R. B., & Hopper, C. H. (1986). Decision making by parole interviewers: The effect of case and interview factors. *Law and Human Behavior*, *10*, 203–214.

Ruby, C. L., & Brigham, J. C. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations. *Psychology, Public Policy, and Law*, *3*, 705–737.

Ruby, C. L., & Brigham, J. C. (1998). Can Criteria-Based Content Analysis distinguish between true and false statements of African-American speakers? *Law and Human Behavior*, *22*, 369–388.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*, 331–349.

Ruva, C., & Bryant, J. B. (1998). *The impact of age, speech style, and question form on perceptions of witness credibility and trial outcome*. Paper presented at the AP-LS Biennial conference, Edondo Beach, CA, March 1998.

Saarni, C. (1979). Children's understanding of display rules for expressive behavior. *Developmental Psychology*, *15*, 424–429.

Saarni, C. (1984). An observational study of children's attempts to monitor their expressive behavior. *Child Development*, *55*, 1504–1513.

Saarni, C., & von Salisch, M. (1993). The socialization of emotional dissemblance. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 106–125). New York: Guilford Press.

Sabbagh, M. A., Moses, L. J., & Shiverick, S. (2006). Executive functioning and preschoolers' understanding of false beliefs, false photographs, and false signs. *Child Development*, *77*, 1034–1049.

Sagarin, B. J., Rhoads, K. L., & Cialdini, R. B. (1998). Deceiver's distrust: Denigration as a consequence of undiscovered deception. *Personality and Social Psychology Bulletin*, *11*, 1167–1176.

Santarcangelo, M., Cribbie, R. A., & Ebesu Hubbard, A. S. (2004). Improving accuracy of veracity judgement through cue training. *Perceptual and Motor Skills*, *98*, 1039–1048.

Santtila, P., Roppola, H., & Niemi, P. (1999). Assessing the truthfulness of witness statements made by children (aged 7–8, 10–11, and 13–14) employing scales derived from Johnson's and Raye's model of Reality Monitoring. *Expert Evidence*, *6*, 273–289.

Santtila, P., Roppola, H, Runtti, M., & Niemi, P. (2000). Assessment of child witness statements using criteria-based content analysis (CBCA): The effects of age, verbal ability, and interviewer's emotional style. *Psychology, Crime, & Law*, *6*, 159–179.

Sapir, A. (1987/2000). *The LSI course on scientific content analysis (SCAN)*. Phoenix, AZ: Laboratory for Scientific Interrogation.

Saxe, L. (1991). Science and the GKT Polygraph: A theoretical critique. *Integrative Physiological and Behavioral Science*, *26*, 223–231.

Saxe, L. (1994). Detection of deception: Polygraph and integrity tests. *Current Directions in Psychological Science*, *3*, 69–73.

Saxe, L., & Ben-Shakhar, G. (1999). Admissibility of polygraph tests: The application of scientific standards post-Daubert. *Psychology, Public Policy and Law*, *5*, 203–223.

Saxe, L., Dougherty, D., & Cross, T. (1985). The validity of polygraph testing: Scientific analysis and public controversy. *American Psychologist*, *40*, 355–366.

Saywitz, K. J. (2002). Developmental underpinnings of children's testimony. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony: A handbook of psychological research and forensic practice* (pp. 3–20). Chichester, England: John Wiley & Sons, Ltd.

Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization and model. *Psychological Bulletin*, *92*, 641–669.

Schneider, S. M., & Kintz, B. L. (1977). The effect of lying upon foot and leg movement. *Bulletin of the Psychonomic Society*, *10*, 451–453.

Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 171–181.

Schul, Y., Burnstein, E., & Bardi, A. (1996). Dealing with deceptions that are difficult to detect: Encoding and judgment as a function of preparing to receive invalide information. *Journal of Experimental Social Psychology*, *32*, 228–253.

Schul, Y., Mayo, R., Burnstein, E., & Yahalom, N. (2007). How people cope with uncertainty due to chance or deception. *Journal of Experimental Social Psychology*, *43*, 91–103.

Schweitzer, M. E., Brodt, S. E., & Croson, R. T. A. (2002). Seeing and believing: Visual access and the strategic use of deception. *The International Journal of Conflict Management*, *13*, 258–275.

Schweitzer, M. E., & Croson, R. (1999). Curtailing deception: The impact of direct questions on lies and omissions. *International Journal of Conflict Management*, *10*, 225–248.

Schweitzer, M. E., Herhsey, J. C., & Bradlow, E. T. (2006). Promises and lies: Restoring violated trust. *Organizational Behavior and Human Decision Processes*, *101*, 1–9.

Seager, P., & Wiseman, R. (1999). Fooling all of the people hald of the time? *Science Spectra*, *15*, 32–37.

Seiter, J. S., Bruschke, J., & Bai, C. (2002). The acceptability of deception as a function of perceivers' culture, deceiver's intention, and deceiver-deceived relationship. *Western Journal of Communication*, *66*, 158–180.

Seto, M. C., Khattar, N. A., Lalumiere, M. L., & Quinsey, V. L. (1997). Deception and sexual strategy in psychopathy. *Personality and Individual Differences*, *22*, 301–307.

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess 'guilty knowledge'. *Journal of Applied Psychology*, *85*, 30–37.

Shallice, T., & Burgess, P. (1994). Supervisory control of action and thought selection. In L. Weiskrantz, A. Baddeley, & D. Alan (Eds.), *Attention: Selection, awareness and control: A tribute to Donald Broadbent* (pp. 171–187). New York: Clarendon Press.

Shennum, W. A., & Bugenthal, D. B. (1982). The development of control over affective expression in nonverbal behavior. In R. S. Feldman (Ed.), *Development of nonverbal behavior in children* (pp. 101–121). New York: Springer-Verlag.

Siegman, A. W., & Reynolds, M. A. (1983). Self-monitoring and speech in feigned and unfeigned lying. *Journal of Personality and Social Psychology*, *45*, 1325–1333.

Silberman, S. (2006). Don't even think about lying: How brain scans are reinventing the science of lie detection. *Wired magazine*, issue 14.01, January.

Sinclair, A. (1996). Young children's practical deceptions and their understanding of false belief. *New Ideas in Psychology*, *4*, 157–173.

Sitton, S. C., & Griffin, S. T. (1981). Detection of deception from clients' eye contact patterns. *Journal of Counseling Psychology*, *28*, 269–271.

Smith, A. (1983). Nonverbal communication among black female dyads: An assessment of intimacy, gender and race. *Journal of Social Issues*, *39*, 55–67.

Smith, H. J., Archer, D., & Costanzo, M. (1991). "Just an hunch": Accuracy and awareness in person perception. *Journal of Nonverbal Behavior*, *15*, 3–19.

Smith, N. (2001). Reading between the lines: An evaluation of the scientific content analysis technique (SCAN). Police research series paper 135. London: UK Home Office, Research, Development and Statistics Directorate.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, *32*, 25–38.

Sodian, B. (1991). The development of deception in young children. *British Journal of Developmental Psychology*, *9*, 173–188.

Sokolov, A. N. (1963). *Perception and the conditioned reflex*. Oxford, England: Pergamon Press.

Spence, S. A. (in press). Playing Devil's Advocate: The case against fMRI lie detection. *Legal and Criminological Psychology*.

Spence, S. A., Farrow, T. F. D., Herford, A. E., Wilkinson, I. D., Zheng, Y., & Woodruff, P. W. R. (2001). Behavioural and functional anatomical correlates of deception in humans. *Neuroreport: For Rapid Communication of Neuroscience Research*, *12*, 2849–2853.

Spence, S. A., Farrow, T. F. D., Leung, D. Shah, S., Reilly, B., Rahman, A., & Herford, A. (2003). Lying as an executive function. In P. W. Halligan, C. Bass, & D. A. Oakley (Eds.), *Malingering and illness deception* (pp. 255–266). Oxford, England: Oxford University Press.

Spence, S. A., Hunter, M. D., Farrow, T. F. D., Green, R. D., Leung, D. H. & Hughes, C. J. (2004). A cognitive neurobiological account of deception:

Evidence from functional neuroimaging. *Philosophical Transactions of the Royal Society of London*, *359*, 1755–1762.

Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, *11*, 373–397.

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

Sporer, S. L., & Schwandt, B. (2006a). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, *20*, 421–446.

Sporer, S. L., & Schwandt, B. (2006b). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law*, *13*, 1–34.

Sporer, S. L., & Sharman, S. J. (2006). Should I believe this? Reality Monitoring of accounts of self-experiences and invented recent and distant autobiographical events. *Applied Cognitive Psychology*, 20, 837–854.

Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, *27*, 76–105.

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

Steller, M., & Boychuk, T. (1992). Childrens as witnesses in sexual abuse cases: Investigative interview and assessment techniques. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp.47–73). New York: John Wiley & Sons, Inc.

Steller, M., Haenert, P., & Eiselt, W. (1987). Extraversion and the detection of information. *Journal of Personality and Social Psychology*, *21*, 334–342.

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

Steller, M., & Wellershaus, P. (1996). Information enhancement and credibility assessment of child statements: The impact of the cognitive interview on criteria-based content analysis. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law, and criminal justice: international developments in research and practice* (pp. 118–127). Berlin: de Gruyter.

Steller, M., Wellershaus, P., & Wolf, T. (1988, June). *Empirical validation of Criteria-Based Content Analysis*. Paper presented at the NATO Advanced Study Institute on Credibility Assessment in Maratea, Italy.

Stephenson, G. M., & Moston, S. J. (1994). Police interrogation. *Psychology, Crime, & Law*, *1*, 151–157.

Stern, J. A., Walrath, L. C., & Goldstein, R. (1984). The endogenous eyeblink. *Psychophysiology, 21*, 22–33.

Sternberg, K. J., Lamb, M. E., Davies, G. M., & Westcott, H. L. (2001). 'Memorandum of good practice: Theory versus application'. *Child Abuse and Neglect*, *25*, 669–681.

Sternberg, K. J., Lamb, M. E., Esplin, P. W., Orbach, Y., & Hershkowitz, I. (2002). Using a structured interview protocol to improve the quality of investigative interviews. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 409–436). Maywah, NJ: Erlbaum.

Sternberg, K. J., Lamb, M. E., Hershkowitz, I., Esplin, P. W., Redlich, A., & Sunshine, N. (1996). The relationship between investigative utterance types and the informativeness of child witnesses. *Journal of Applied Developmental Psychology*, *17*, 439–451.

Stiff, J. B., Corman, S., Krizek, B., & Snider, E. (1994). Individual differences and changes in nonverbal behavior: Unmasking the changing faces of deception. *Communication Research*, *21*, 555–581.

Stiff, J. B., Hale, J. L., Garlick, R., & Rogan, R. (1990). Effect of cue incongruence and social normative influences on individual judgments of honesty and deceit. *The Southern Communication Journal*, *55*, 206–229.

Stiff, J. B., Kim, H. J., & Ramesh, C. N. (1992). Truth biases and aroused suspicion in relational deception. *Communication Research*, *19*, 326–345.

Stiff, J. B., & Miller, G. R. (1986). "Come to think of it . . . ": Interrogative probes, deceptive communication, and deception detection. *Human Communication Research*, *12*, 339–357.

Stiff, J. B., Miller, G. R., Sleight, C., Mongeau, P., Garlick, R., & Rogan, R. (1989). Explanations for visual cue primacy in judgments of honesty and deceit. *Journal of Personality and Social Psychology*, *56*, 555–564.

Stouthamer-Loeber, M. (1986). Lying as a problem behavior in children: A review. *Clinical Psychology Review*, *6*, 267–289.

Streeter, L. A., Krauss, R. M., Geller, V., Olson, C., & Apple, W. (1977). Pitch changes during attempted deception. *Journal of Personality and Social Psychology*, *24*, 12–21.

Strömwall, L. A., Bengtsson, L., Leander, L., & Granhag, P. A. (2004). Assessing children's statements: The impact of a repeated experience on CBCA and RM ratings. *Applied Cognitive Psychology*, *18*, 653–668.

Strömwall, L. A., & Granhag, P. A. (2003a). Affecting the perception of verbal cues to deception. *Applied Cognitive Psychology*, *17*, 35–49.

Strömwall, L. A., & Granhag, P. A. (2003b). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime, & Law*, *9*, 19–36.

Strömwall, L. A., & Granhag, P. A. (2005). Children's repeated lies and truths: Effects on adults' judgments and reality monitoring scores. *Psychiatry, Psychology, & Law*, *12*, 345–356.

Strömwall, L. A., & Granhag, P. A. (2007). Detecting deceit in pairs of children. *Journal of Applied Social Psychology*, 37, 1285–1304.

Strömwall. L. A., Granhag, P. A., & Hartwig, M. (2004). Practitioners' beliefs about deception. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 229–250). Cambridge, England: Cambridge University Press.

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

Strömwall, L. A, Granhag, P. A., & Landström, S. (2007). Children's prepared and unprepared lies: Can adults see through their strategies? *Applied Cognitive Psychology*, *21*, 457–471.

Strömwall, L. A., Hartwig, M., & Granhag, P. A. (2006). To act truthfully: Nonverbal behaviour and strategies during a police interrogation. *Psychology, Crime, & Law*, *12*, 207–219.

Svebak, S. (2006). *Polygraph tests in Norwegion courts: Legal statud and future potential*. Paper presented at the European Expert Meeting on Polygraph Testing. University of Maastricht, Maastricht, the Netherlands, 29–31 March.

Talwar, V., & Lee. K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioural Development*, *26*, 437–444.

Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2006). Adults' judgements of children's coached reports. *Law and Human Behavior*, *30*, 561–570.

Talwar, V., Murphy, S. M., & Lee, K. (2007). White lie-telling in children for politeness purposes. *International Journal of Behavioural Development*, *31*, 1–11.

Tate, C. S., Warren, A. R., & Hess, T. H. (1992). Adults' liability for children's 'Lie-Ability': Can adults coach children to lie successfully? In S. J. Ceci, M. DeSimone Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 69–88). Hillsdale, NJ: Erlbaum.

Taylor, R., & Hick, R. F. (2007). Believed cues to deception: Judgements in self-generated serious and trivial situations. *Legal and Criminological Psychology, 12,* 321–332.

Taylor, R., & Hill-Davies, C. (2004). Parents' and non-parents' beliefs about the cues to deception in children. *Psychology, Crime, & Law*, *10*, 455–464.

Taylor, R., & Vrij, A. (2001). The effects of varying stake and cognitive complexity on beliefs about the cues to deception. *International Journal of Police Science and Management*, *3*, 111–123.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 1157–1172.

Tecce, J. J. (1992). Psychology, physiology and experimental. *In McGraw-Hill yearbook of science and technology* (pp. 375–377). New York: McGraw-Hill.

Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (volume 11, pp. 288–338). New York: Academic Press.

The Global Deception Team (2006). A world of lies. *Journal of Cross-Cultural Psychology*, *37*, 60–74.

Tickle-Degnen, L. (2006). Nonverbal behavior and its functions in the ecosystem of rapport. In V. Manusov & M. L. Patterson (Eds.), *The SAGE handbook of nonverbal communication* (pp. 381–399). Thousand Oaks, CA: Sage.

Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychologival Inquiry*, *1*, 285–293.

Toglia, M. P., Read, J. D, Ross, D. F., & Lindsay, R. C. L. (Eds.) (2006). *The handbook of eyewitness psychology: Volume One: Memory for events.* Mahwah, NJ: Erlbaum.

Tooke, W., & Camire, L. (1991). Patterns of deception in intersexual and intrasexual mating strategies. *Ethology and Sociobiology*, *12,* 345–364.

Toris, C., & DePaulo, B. M. (1984). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology*, *47*, 1063–1073.

Trankell, A. (1963). *Vittnespsykologins Arbetsmetoder*. Stockholm, Sweden: Liber.

Trankell, A. (1972). *Reliability of evidence*. Stockholm, Sweden: Beckmans.

Tsiamyrtzis, P., Dowdall, J., Shastri, D., Pavlidis, I. T., Frank, M. G., & Ekman, P. (2007). Imaging facial physiology for the detection of deceit. *International Journal of Computer Vision*, *71*, 197–214.

Tully, B. (1998). Reliability of criteria-based content analysis of child witness statements: Cohen's kappa doesn't matter. *Legal and Criminological Psychology*, *3*, 183–188.

Tully, B. (1999). Statement validation. In D. Canter & L. Alison (Eds.), *Interviewing and deception* (pp. 83–104). Aldershot, England: Darmouth.

Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology*, *11*, 69–89.

Tye, M. C., Amato, S. L., Honts, C. R., Kevitt, M. K., & Peters, D. (1999). The willingness of children to lie and the assessment of credibility in an ecologically relevant laboratory setting. *Applied Developmental Science*, *3*, 92–109.

Tyler, J. M., Feldman, R. S., & Reichert, A. (2006). The price of deceptive behavior: Disliking and lying to people who lie to us. *Journal of Experimental Social Psychology*, *42*, 69–77.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1134.

Undeutsch, U. (1967). Beurteilung der Glaubhaftigkeit von Aussagen. In U. Undeutsch (Ed.), *Handbuch der Psychologie Vol. 11: Forensische Psychologie* (pp. 26–181). Göttingen, Germany: Hogrefe.

Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 27–56). Deventer, The Netherlands: Kluwer.

Undeutsch, U. (1984). Courtroom evaluation of eyewitness testimony. *International Review of Applied Psychology*, *33*, 51–67.

Undeutsch, U. (1989). The development of statement reality analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 101–121). Dordrecht, The Netherlands: Kluwer.

Van den Adel, H. M. (1997). *Handleiding verdachtenverhoor*. Den Haag: VUGA-Uitgeverij.

Van der Plas, M. (2006). *The use of the polygraph in Belgium*. Paper presented at the European Expert Meeting on Polygraph Testing. University of Maastricht, Maastricht, the Netherlands, 29–31 March.

Van Rossum, W. (1998). *Verschijnen voor de rechter: Hoe het hoort en het ritueel van Turkse verdachten in de rechtszaal*. Amsterdam, The Netherlands: Uitgeverij duizend en een.

Vendemia, J. M. C., Buzan, R. F., & Green, E. P. (2005a). Practice effects, workload, and reaction time in deception. *Americal Journal of Psychology*, *118*, 413–429.

Vendemia, J. M. C., Buzan, R. F., & Simon-Dack, S. L. (2005b). Reaction time of motor responses in two-stimulus paradigms involving deception and congruity with varying levels of difficulty. *Behavioural Neurology*, *16*, 25–36.

Verschuere, B., Crombez, G., De Clercq, A., & Koster, E. H. W. (2004a). Autonamic and behavioural responding to concealed information: Differentiating orienting and defensive responses. *Psychophysiology*, *41*, 461–466.

Verschuere, B., Crombez, G., & Koster, E. H. W. (2004b). Orienting to guilty knowledge. *Cognition and Emotion*, *18*, 265–279.

Verschuere, B., Crombez, G., de Clercq, A., & Koster, E. H. W. (2005). Psychopathic traits and autonomic responding to concealed information in a prison sample. *Psychophysiology*, *42*, 239–245.

Volbert, R., & Van der Zanden, R. (1996). Sexual knowledge and behavior of children up to 12 years: What is age-appropriate? In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law, and criminal justice: international developments in research and practice* (pp. 198–216). Berlin: de Gruyter.

Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills and physical characteristics on impression formation. *Journal of Social Psychology*, *133*, 601–611.
Vrij, A. (1994). The impact of information and setting on detection of deception by police detectives. *Journal of Nonverbal Behavior*, *18*, 117–137.
Vrij, A. (1995). Behavioral correlates of deception in a simulated police interview. *Journal of Psychology: Interdisciplinary and Applied*, *129*, 15–29.
Vrij, A. (1997). Wearing black clothes: The impact of offenders' and suspects' clothing on impression formation. *Applied Cognitive Psychology*, *11*, 47–53.
Vrij, A. (1998). Telling and detecting lies: Some future directions in research. *Forensic Update*, 54, 14–19.
Vrij, A. (2000). Telling and detecting lies as a function of raising the stakes. In C. M. Breur, M. M. Kommer, J. F. Nijboer, & J. M. Reintjes (Eds.), *New trends in criminal investigation and evidence, volume 2* (pp. 699–709). Antwerpen, Belgium: Intersentia.
Vrij, A. (2001). Implicit lie detection. *The Psychologist*, *14*, 58–60.
Vrij, A. (2002a). Deception in children: A literature review and implications for children's testimony. In H. L. Westcott, G. M. Davies, & R. H. C. Bull (Eds.), *Children's testimony* (pp. 175–194). Chichester: John Wiley and Sons, Ltd.
Vrij, A. (2002b). Telling and detecting lies. In N. Brace & H. L. Westcott (Eds.), *Applying psychology* (pp. 179–241). Milton Keynes: Open University.
Vrij, A. (2003). We will protect your wife and child, but only if you confess: Police interrogations in England and the Netherlands. In P. J. van Koppen & S. D. Penrod (Eds.), *Adversarial versus inquisitorial justice: Psychological perspectives on criminal justice systems* (pp. 55–79). New York: Plenum.
Vrij, A. (2004a). Guidelines to catch a liar. In P. A. Granhag & L. A. Strömwall (Eds.), *Deception detection in forensic contexts* (pp. 287–314). Cambridge, England: Cambridge University Press.
Vrij, A. (2004b). 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. (2006a). Challenging interviewees during interviews: The potential effects on lie detection. *Psychology, Crime, & Law*, *12*, 193–206.
Vrij, A. (2006b). Nonverbal communication and deception. In V. Manusov & M. L. Patterson (Eds.), *The Sage handbook of nonverbal communication* (pp. 341–359). Thousand Oaks, CA: Sage.
Vrij, A. (2007). Deception: A social lubricant and a selfish act. In K. Fiedler (Ed.), *Frontiers of social psychology: Social communication* (pp. 309–342). New York: Psychology Press.
Vrij, A. (in press). Nonverbal dominance versus verbal accuracy in lie detection: A plea to change police practice. *Criminal Justice and Behavior*.
Vrij, A., & Akehurst, L. (1997). The existence of a black clothing stereotype: The impact of a victim's black clothing on impression formation. *Psychology, Crime, & Law*, *3*, 227–237.
Vrij, A., Akehurst, L., Brown, L., & Mann, L. (2006). Detecting lies in young children, adolescents and adults. *Applied Cognitive Psychology*, *20*, 1225–1237.
Vrij, A., Akehurst, L., & Knight, S. (2006). Police officers', social workers', teachers' and the general public's beliefs about deception in children, adolescents and adults. *Legal and Criminological Psychology, 11,* 297–312.

Vrij, A., Akehurst, L., & Morris, P. M. (1997). Individual differences in hand movements during deception. *Journal of Nonverbal Behavior*, *21*, 87–102.
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 Behaviour*, *26*, 261–283.
Vrij, A., Akehurst, L. Soukara, S., & Bull, R. (2004a). Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. *Human Communication Research*, *30*, 8–41.
Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004b). Let me inform you how to tell a convincing story: CBCA and Reality Monitoring scores as a function of age, coaching and deception. *Canadian Journal of Behavioural Science*, *36*, 113–126.
Vrij, A., Akehurst, L., Van Dalen, D., Van Wijngaarden, J. J., & Foppes, J. H. (1996). Nonverbaal gedrag en misleiding. *Tijdschrift voor de Politie*, *58*, 11–14.
Vrij, A., & Baxter, M. (1999). Accuracy and confidence in detecting truths and lies in elaborations and denials: Truth bias, lie bias and individual differences. *Expert Evidence*, *7*, 25–36.
Vrij, A., Dragt, A. W., & Koppelaar, L. (1992b). Interviews with ethnic interviewees: Nonverbal communication errors in impression formation. *Journal of Community and Applied Social Psychology*, *2*, 199–209.
Vrij, A., Edward, K., & Bull, R. (2001a). People's insight into their own behaviour and speech content while lying. *British Journal of Psychology*, *92*, 373–389.
Vrij, A., Edward, K., & Bull, R. (2001b). Police officers' ability to detect deceit: The benefit of indirect deception detection measures. *Legal and Criminological Psychology*, *6*, 2, 185–197.
Vrij, A., Edward, K., & Bull, R. (2001c). Stereotypical verbal and nonverbal responses while deceiving others. *Personality and Social Psychology Bulletin*, *27*, 899–909.
Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, *24*, 239–263.
Vrij, A., Ennis, E., Farman, S., & Mann, S. (2006). *People's perceptions of their truthful and deceptive interactions in daily life*. Manuscript submitted for publication.
Vrij, A., Evans, H., Akehurst, L., & Mann, S. (2004). Rapid judgements in assessing verbal and nonverbal cues: Their potential for deception researchers and lie detection. *Applied Cognitive Psychology*, *18*, 283–296.
Vrij, A., & Fischer, A. (1995). The expression of emotions in simulated rape interviews. *Journal of Police and Criminal Psychology*, *10*, 64–67.
Vrij, A., & Fischer, A. (1997). The role of displays of emotions and ethnicity in judgements of rape victims. *International Review of Victimology*, *4*, 255–265.
Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences, 10*, 141–142.
Vrij, A., Fisher, R., Mann, S., & Leal, S. (in press). Increasing cognitive load in interviews to detect deceit. In B. Milne, S. Savage, & T. Williamson (Eds.), *International developments in investigative interviewing*. Uffculme: Willan Publishing.
Vrij, A., Floyd, M., & Ennis, E. (2003). Telling lies to strangers or close friends: Its relationship with attachment style. In S. P. Shohov (Ed.), *Advances*

in psychological research, volume 20 (pp. 61–74). New York: NovaScience Publishers.

Vrij, A., Foppes, J. H., Volger, D. M., & Winkel, F. W. (1992). Moeilijk te bepalen wie de waarheid spreekt: Non-verbaal gedrag belangrijkste indicator. *Algemeen Politie Blad*, *141*, 13–15.

Vrij, A., & Graham, S. (1997). Individual differences between liars and the ability to detect lies. *Expert Evidence*, *5*, 144–148.

Vrij, A., & Granhag, P. A. (2007). Interviewing to detect deception. In S. A. Christianson (Ed.), *Offenders' memories of violent crimes* (pp. 279–304). Chichester, England: John Wiley & Sons, Ltd.

Vrij, A., Granhag, P. A., & Mann, S. (in press). Good liars. *Review of Policy Research*.

Vrij, A., Harden, F., Terry, J., Edward, K., & Bull, R. (2001). The influence of personal characteristics, stakes and lie complexity on the accuracy and confidence to detect deceit. In R. Roesch, R. R. Corrado, & R. J. Dempster (Eds.), *Psychology in the courts: International advances in knowledge* (pp. 289–304). London: Routlegde.

Vrij, A., & Heaven, S. (1999). Vocal and verbal indicators of deception as a function of lie complexity. *Psychology, Crime, & Law*, *5*, 203–315.

Vrij, A., & Holland, M. (1998). Individual differences in persistence in lying and experiences while deceiving. *Communication Research Reports*, *15*, 299–308.

Vrij, A., Kneller, W., & Mann, S. (2000). The effect of informing liars about criteria-based content analysis on their ability to deceive CBCA-raters. *Legal and Criminological Psychology*, *5*, 57–70.

Vrij, A., & Lochun, S. (1997). Neuro-linguïstisch verhoren. In P. J. Van Koppen, D. J. Hessing, & H. F. M. Crombag (Eds.), *Het hart van de zaak: Psychologie van het recht* (pp. 493–505). Deventer, the Netherlands: Kluwer.

Vrij, A., & Mann, S. (2001a). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology*, *15*, 187–203.

Vrij, A., & Mann, S. (2001b). Who killed my relative? Police officers' ability to detect real-life high-stake lies. *Psychology, Crime, & Law*, *7*, 119–132.

Vrij, A., & Mann, S. (2003a). Deceptive responses and detecting deceit. In P. W. Halligan, C. Bass, & D. Oakley (Eds.), *Malingering and illness deception: Clinical and theoretical perspectives* (pp. 348–362). Oxford, England: Oxford University Press.

Vrij, A., & Mann, S. (2003b). Telling and detecting true lies: Investigating and detecting the lies of murderers and thieves during police interviews. In M. Verhallen, G. Verkaeke, P. J. van Koppen, & J. Goethals (Eds.), *Much ado about crime: Chapters on psychology and law* (pp. 185–208). Brussels, Belgium: Uitgeverij Politeia.

Vrij, A., & Mann, S. (2004). Detecting deception: The benefit of looking at a combination of behavioral, auditory and speech content related cues in a systematic manner. *Group Decision and Negotiation*, *13*, 61–79.

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

Vrij, A., & Mann, S. (2006). Criteria-Based Content Analysis: An empirical test of its underlying processes. *Psychology, Crime, & Law*, *12*, 337–349.

Vrij, A., Mann, S., & Fisher, R. (2006a). An empirical test of the Behaviour Analysis Interview. *Law and Human Behavior, 30,* 329–345.

Vrij, A., Mann, S., & Fisher, R. (2006b). Information-gathering vs accusatory interview style: Individual differences in respondents' experiences. *Personality and Individual Differences*, *41*, 589–599.

Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., & Bull, R. (in press). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*.

Vrij, A., Mann, S., Kristen, S., & Fisher, R. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior, 31*, 499–518.

Vrij, A., Mann, S., Robbins, E., & Robinson, M. (2006). Police officers ability to detect deception in high-stakes situations and in repeated lie detection tests. *Applied Cognitive Psychology*, *20*, 741–755.

Vrij, A., Nunkoosing, K., Paterson, B., Oosterwegel, A., & Soukara, S. (2002). Characteristics of secrets and the frequency, reasons and effects of secrets keeping and disclosure. *Journal of Community and Applied Social Psychology*, *12*, 56–70.

Vrij, A., Pannell, H., & Ost, J. (2005). The influence of social pressure and black clothing on crime judgements. *Psychology, Crime, & Law*, *11*, 265–274.

Vrij, A., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior*, *20*, 65–80.

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

Vrij, A., & Taylor, R. (2001). Police officers' and students' beliefs about telling and detecting little and serious lies. *International Journal of Police Science & Management*, *5*, 1–9.

Vrij, A., & Van Wijngaarden, J. J. (1994). Will truth come out? Two studies about the detection of false statements expressed by children. *Expert Evidence*, *3*, 78–84.

Vrij, A., & Winkel, F. W. (1991). Cultural patterns in Dutch and Surinam nonverbal behavior: An analysis of simulated police/citizen encounters. *Journal of Nonverbal Behavior*, *15*, 169–184.

Vrij, A., & Winkel, F. W. (1992a). Crosscultural police-citizen interactions: The influence of race, beliefs and nonverbal communication on impression formation. *Journal of Applied Social Psychology*, *22*, 1546–1559.

Vrij, A., & Winkel, F. W. (1992b). Social skills, distorted perception and being suspect: Studies in impression formation and the ability to deceive. *Journal of Police and Criminal Psychology*, *8*, 2–6.

Vrij, A., & Winkel, F. W. (1994). Perceptual distortions in crosscultural interrogations: The impact of skin color, accent, speech style and spoken fluency on impression formation. *Journal of Cross-Cultural Psychology*, *25*, 284–296.

Vrij, A. & Winkel, F. W. (1995). Detection of false statements in first and third graders: The development of a nonverbal detection instrument. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law and criminal justice: International developments in research and practice* (p. 221–230). Berlin: de Gruyter.

Vrij, A., Winkel, F. W., Akehurst, L., (1997). Police officers' incorrect beliefs about nonverbal indicators of deception and its consequences. In J. F. Nijboer & J. M. Reijntjes (Eds.), *Proceedings of the first world conference on new trends in criminal investigation and evidence* (pp. 221–238). Lelystad, the Netherlands: Koninklijke Vermande.

Vrij, A., Winkel, F. W., & Koppelaar, L. (1988). Culturele verschillen in nonverbaal gedrag: De persoonlijke ruimte van Nederlanders en Surinamers. *Migrantenstudies*, *4*, 40–49.

Vrij, A., Winkel, F. W., & Koppelaar, L. (1991). Interactie tussen politiefunctionarissen en allochtone burgers: twee studies naar de frequentie en het effect van aan- en wegkijken op de impressieformatie. *Nederlands Tijdschrift voor de Psychologie*, *46*, 8–20.

Vrij, J. (2001). Verzet tegen angst. *Kontakt door Aantreden*, *56*, 4.

Wagenaar, W. A., & Dalderop, A. (1994). *Remembering the zoo: A comparison of true and false stories told by pairs of witnesses*. Unpublished manuscript, Department of Experimental Psychology, Leiden University, the Netherlands.

Wagenaar, W. A., & Groeneweg, J. (1990). The memory of concentration camp survivors. *Applied Cognitive Psychology*, *4*, 77–87.

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, *17*, 755–744.

Walczyk, J. J., Schwartz, J. P., Clifton, R., Adams, B., Wei, M., & Zha, P. (2005). Lying person-to-person about live events: A cognitive framework for lie detection. *Personnel Psychology*, *58*, 141–170.

Walker, A. G., & Warren, A. R. (1995). The language of the child abuse interview: Asking the questions, understanding the answers. In T. Ney (Ed.), *True and false allegations in child sexual abuse: Assessment and case management* (pp. 153–162). New York: Brunner-Mazel.

Walkley, J. (1985). Reading the suspect. *Police Review*, 15 February.

Wallbott, H. G., & Scherer, K. R. (1991). Stress specifics: Differential effects of coping style, gender, and type of stressor on automatic arousal, facial expression, and subjective feeling. *Journal of Personality and Social Psychology*, *61*, 147–156.

Walters, S. B. (1996). *Kinesic interview and interrogation*. Boca Raton, Florida: CRC Press.

Waltman, J. L. (1983). Nonverbal communication in interrogation: Some applications. *Journal of Police and Science Administration*, *11*, 166–169.

Wan Cheng, K. H., & Broadhurst, R. (2005). The detection of deception: The effects of first and second language on lie detection ability. *Psychiatry, Psychology and Law*, *12*, 107–118.

Wang, G., Chen, H., & Atabakhsh, H. (2004). Criminal identity deception and deception detection in law enforcement. *Group Decision and Negotitation*, *13*, 111–127.

Watson, D. C., & Sinka, B. K. (1993). Individual differences, social arousal and the electrodermal detection of deception. *Personality and Individual Differences*, *15*, 75–80.

Wegener, H. (1989). The present state of statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 121–134). Dordrecht, the Netherlands: Kluwer.

Weiss, B., & Feldman, R. S. (2006). Looking good and lying to do it: Deception as an impression management strategy in job interviews. *Journal of Applied Social Psychology*, *36*, 1070–1086.

Wells, G. L., & Leippe, M. R. (1981). How do triers of fact infer accuracy of eyewitness identification? Using memory of peripheral details can be misleading. *Journal of Applied Psychology*, *66*, 682–687.

Wells, G. L., & Loftus, E. F. (1991). Commentary: Is this child fabricating? Reactions to a new assessment technique. In J. Doris (Ed.), *The suggestibility* *of children's recollections* (pp. 168–171). Washington, DC: American Psychological Association.

Wessel, E., Drevland, G., Eilertsen, D. E., & Magnussen, S. (2006). Credibility of the emotional witness: A comparison of ratings by laypersons and court judges. *Law and Human Behavior, 30,* 221–230.

Westcott, H., & Brace, N. (2002). Psychological factors in witness evidence and identification. In H. Westcott & N. Brace (2002). *Applying psychology* (pp. 117–178). Milton Keynes, England: Open University Press.

Westcott, H. L., Davies, G. M., & Clifford, B. R. (1991). Adults' perceptions of children's videotaped truthful and deceptive statements. *Applied Social Psychology*, *31*, 2322–2338.

Westcott, H. L., & Kynan, S. (2004). The application of a 'story-telling' framework to investigative interviews for suspected child sexual abuse. *Legal and Criminological Psychology*, *9*, 37–56.

Westcott, H. L., & Kynan, S. (2006). Interviewer practice in investigative interviews for suspected child sexual abuse. *Psychology, Crime, & Law*, *12*, 367–382.

Westcott, H. L., Kynan, S., & Few, C. (2005). Improving the quality of investigative interviews for suspected child abuse: A case study. *Psychology, Crime, & Law*, *11*, 77–96.

White, C. H., & Burgoon, J. K. (2001). Adaptation and communicative design: Patterns of interaction in truthful and deceptive conversations. *Human Communication Research*, *27*, 9–37.

Whitty, M. T. (2002). Liar, liar! An examination of how open, supportive and honest people are in chat rooms. *Computers in Human Behavior*, *18*, 343–352.

Wilcox, D. T., Sosnowski, D., & Middleton, D. (2000). Polygraphy and sex offenders. *Forensic Update*, *61*, 20–25.

Wild, J. (2005). Brain imaging ready to detect terrorists, say neuroscientists. *Nature*, *437*, 457.

Williams, S. (2001). Sexual lying among college students in close and casual relationships. *Journal of Applied Social Psychology*, *31*, 2322–2338.

Williamson, T. (1993). From interrogation to investigative interviewing: Strategic trends in police questioning. *Journal of Community and Applied Social Psychology*, *3*, 89–99.

Wilson, A. E., Smith, M. D., & Ross, H. S. (2003). The nature and effects of young children's lies. *Social Development, 12*, 21–45.

Wilson, D. S., Near, D. C., & Miller, R. R. (1998). Individual differences in Machiavellianism as a mix of cooperative and exploitative strategies. *Evolution and Human Behavior*, *19*, 203–212.

Winkel, F. W., & Koppelaar, L. (1991). Rape victims' style of self-presentation and secondary victimization by the environment. *Journal of Interpersonal Violence*, *6*, 29–40.

Winkel, F. W., & Vrij, A. (1990). Interaction and impression formation in a cross-cultural dyad: Frequency and meaning of culturally determined gaze behaviour in a police interview setting. *Social Behaviour*, *5*, 335–350.

Winkel, F. W., & Vrij, A. (1995). Verklaringen van kinderen in interviews: Een experimenteel onderzoek naar de diagnostische waarde van Criteria Based Content Analysis. *Tijdschrift voor Ontwikkelingspsychologie*, *22*, 61–74.

Winkel, F. W., Vrij, A., Koppelaar, L., & Van der Steen, J. (1991). Reducing secondary victimisation risks and skilled police intervention: Enhancing the quality of police rape victim encounters through trainingprogrammes. *Journal of Police and Criminal Psychology*, *7*, 2–11.

Wolpe, P. R., Foster K. R., & Langleben, D. D. (2005). Emerging neurotechnologies for lie-detection: Promises and perils. *The American Journal of Bioethics*, *5*, 39–49.
Woodall, W. G., & Burgoon, J. K. (1983). Talking fast and changing attitudes: A critique and clarification. *Journal of Nonverbal Behavior*, *8*, 126–143.
Yang, Y., Raine, A., Lencz, T., Bihrle, S., Lacasse, L., & Coleti, P. (2005). Prefrontal white matter in pathological liars. *British Journal of Psychiatry*, *187*, 320–325.
Yeschke, C.L. (1997). *The art of investigative interviewing*. Newton, Massachusetts: Butterworth-Heinemann
Yuille, J. C. (1988a). *A simulation study of criteria-based content analysis*. Paper presented at the NATO advanced study institute on credibility assessment, Maratea, Italy.
Yuille, J. C. (1988b). The systematic assessment of children's testimony. *Canadian Psychology*, *29*, 247–262.
Yuille, J. C., & Cutshall, J. (1989). Analysis of statements of victims, witnesses, and suspects. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 175–191). Dordrecht, The Netherlands: Kluwer.
Zajonc, R. B. (1980). Compresence. In P. B. Paulus (Ed.), *Psychology of group influence* (pp. 35–60). Hillsdale, NJ: Erlbaum.
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.
Zaragoza, M. S., Payment, K. E., Ackil, J. K., Drivdahl, S. B., & Beck, M. (2001). Interviewing witnesses: Forced confabulation and confirmatory feedback increase false memories. *Psychological Science*, *12*, 473–477.
Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). 'Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, *22*, 1258–1269.
Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004a). Automating linguistics-based cues for detecting deception in text-based asynchronous computer mediated communication. *Group Decision and Negotiation*, *13*, 81–106.
Zhou, L., Burgoon, J. K., Twitchell, D. P., Qin, T., & Nunamaker, J. F. (2004). A comparison of classification models for predicting deception in computer-mediated communication. *Journal of Management Information Systems*, *20*, 139–165.
Zhou, L., Burgoon, J. K., Zhang, D., & Nunamaker, J. F. (2004b). Language dominance in interpersonal deception in computer-mediated communication. *Computers in Human Behavior*, *20*, 381–402.
Zhou, L., & Zang, D. (2006). A comparison of deception behavior in dyad and triadic group decision making in synchronous computer-mediated communication. *Small Group Research*, *37*, 140–164.
Zraick, R. I., Gentry, M. A., Smith-Olinde, L., & Gregg, B. A. (2006). The effect of speaking context on elicitation of habitual pitch. *Journal of Voice*, *20*, 545–554.
Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrance, D. T., & Rosenthal, R. (1979). Facial and vocal cues of deception and honesty. *Journal of Experimental Social Psychology*, *15*, 378–396.
Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981a). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in*

experimental social psychology, volume 14 (pp. 1–57). New York: Academic Press.

Zuckerman, M., Driver, R. E., & Guadagno, N. S. (1985). Effects of segmentation patterns on the perception of deception. *Journal of Nonverbal Behavior*, *9*, 160–168.

Zuckerman, M., Driver, R., & Koestner, R. (1982). Discrepancy as a cue to actual and perceived deception. *Journal of Nonverbal Behavior*, *7*, 95–100.

Zuckerman, M., Koestner, R., & Alton, A. O. (1984). Learning to detect deception. *Journal of Personality and Social Psychology*, *46*, 519–528.

Zuckerman, M., Koestner, R., Colella, M. J. (1985). Learning to detect deception from three communication channels. *Journal of Nonverbal Behavior*, *9*, 188–194.

Zuckerman, M., Koestner, R., & Driver, R. (1981b). Beliefs about cues associated with deception. *Journal of Nonverbal Behavior*, *6*, 105–114.

Zuckerman, M., Speigel, N. H., DePaulo, B. M., & Rosenthal, R. (1982). Nonverbal strategies for decoding deception. *Journal of Nonverbal Behavior*, *6*, 171–187.

Zulawski, D. E., & Wicklander, D. E. (1993). *Practical aspects of interview and interrogation*. Boca Raton, Florida: CRC Press.

Index