The polygraph and the detection of deception

Ewout H. Meijer¹ & Bruno Verschuere²

1 Faculty of Psychology and Neuroscience, Maastricht University, PO Box 616, 6200

MD Maastricht The Netherlands

2 Faculty of Psychology, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium

Corresponding author:

Ewout Meijer

Faculty of Psychology and Neuroscience

Maastricht University

PO Box 616

6200 MD Maastricht

Phone: +31 433884505

Fax: +31 433884196

Email: eh.meijer@psychology.unimaas.nl

Over the last decade, Europe has seen a marked increase in the use of the polygraph for the detection of deception. Belgium and Finland, nowadays regularly use polygraph tests in criminal investigations, and the United Kingdom and the Netherlands have adopted its use in the treatment and monitoring of sex offenders (Meijer & van Koppen, 2008). Still the use of the polygraph for the detection of deception has been debated in the scientific literature for ages. In this contribution we highlight the promises and perils of the use of the polygraph for the detection.

The polygraph

The words 'polygraph' and 'lie detector' are often used synonymously. Technically, the term 'polygraph' refers only to the recording device that is used for registering different physiological parameters. Polygraphs that are specifically designed for lie detection tests used to be briefcase-sized machines that registered physiological signals from the subject by recording them with multiple pens on a lengthy roll of paper. Nowadays, a small amplifier/digitizer and a laptop recording the signals are used in most cases. The sensors attached to the subject are generally (1) two expendable bands positioned around the thorax and the abdomen measuring respiration, (2) two electrodes attached to the inside of the hand measuring electrodermal activity and (3) an inflatable cuff positioned around the upper arm registering blood pressure. These physiological parameters co-vary with a number of psychological processes, including attention and emotion. Consequently, inferences about psychological processes can be made from polygraph recordings. And although it is widely accepted that there is no unique physiological pattern associated with lying, polygraph recording can still be used to infer guilt or innocence.

Detection of deception with the polygraph

The most widely used polygraph technique is known as the Control Question Test (CQT; Reid, 1947). In this type of test, the suspect answers several questions while connected to the polygraph. Among these questions are relevant and control questions. The relevant questions specifically deal with the incident under investigation, for example "on the 25th of March, did you hold up the post office in Chorley?" The control questions have a more generic nature, but also deal with undesirable behavior, for example "In the first 25 years of your life, have you ever done anything illegal?" It is reasoned that the most threatening questions will elicit the strongest physiological responses. The crucial assumption is that the relevant questions are most threatening for guilty suspects, whereas the control questions form the biggest threat to innocent suspects. This latter assumption requires some explanation. A CQT is preceded by a lengthy interview that serves to convince the suspect that the polygraph can determine to an extremely high degree of accuracy determine whether the suspect is lying or not. This means that an innocent examinee can confidently and honestly answer "no" to the relevant questions. After all, the polygraph will show that this is the truth. Furthermore, the interview serves to maneuver the examinee into answering "no" to the control questions by suggesting that confessing illegal activities will negatively influence the test outcome (See box 1). As a result, the innocent examinee is assumed to show the strongest physiological responses to the control questions, fearing that his/her deceptive answer to this question will get him/her convicted for the crime under investigation.

Box 1 approximately here

The CQT has been criticized on several grounds, but it is especially the assumption that innocent suspects will be most concerned about the control questions that is controversial. After all, the stress inducing effect is not a feature of the control question itself, but rather a consequence of how its function is explained to the suspect. As a consequence, the accuracy of the test depends largely on the skills of the polygrapher, rather than on the test itself. According to proponents of the CQT, a skilled polygrapher is capable of formulating control questions and creating an atmosphere in which an innocent examinee will be more worried about the control questions than about the relevant questions (Raskin & Honts, 2002). CQT critics argue that this assumption has no grounding in psychological or psychophysiological research, nor is it convincing in its inner logic (Ben-Shakhar, 2008; Fiedler, Schmid, & Stahl, 2002; Iacono, 2008; Lykken, 1998).

Accuracy

To what extent the problems with the CQT result in erroneous outcomes remains under dispute. Part of this dispute traces back to the difficulties that characterize this type of research. In a typical laboratory study, half of the participants are instructed to commit some kind of mock crime, while the other half remains innocent. Subsequently, all participants are tested with a CQT and the test outcome is determined for each examinee. The problem with such a laboratory study is that participants do not face severe consequences when failing the test, and these studies therefore lack ecological validity. Field studies have more ecological validity, but are plagued by other problems. Most importantly, they lack an objective criterion of guilt or innocence (i.e., ground truth). Often, a criterion such as a conviction or a confession is used to determine the error rate in real life polygraph tests. Both criteria are suboptimal given that they may, directly or indirectly, be influenced by the results of the CQT. Especially the use of confessions made by suspects when confronted with a deceptive test outcome result in a sampling bias overestimating the validity of the CQT (Iacono, 1991).

Keeping these limitations in mind, accuracy estimates of the CQT range from 74% to 89% for guilty examinees, with 1% to 13% false negatives, and 59% to 83% for innocent examinees, with a false positive ratio varying from 10% to 23% (see Table 1).¹ Most recently, a panel of 14 leading American scientists reviewed the literature on the accuracy of the CQT. The 37 laboratory studies and 7 field studies that passed the minimum standards for review showed an accuracy index (area under the ROC curve) of .85 and .89, respectively.² It led the panel to conclude that specific-incident polygraph tests can discriminate lying from truth telling at rates well above chance, though well below perfection (National Research Council, 2003).³

The above allows for three important conclusions. First, the CQT performs above chance level. Second, these figures highlight that the error rate of the CQT can be substantial. Third, with the cut-off points used in practice, the test is especially prone to false positive outcomes. This is problematic for application in the legal arena, as it is alien to legal doctrine abbreviated in the so-called Blackstone Maxim: "Better that ten

¹ These percentages do not necessarily add up to 100% because of an inconclusive category. This inconclusive outcome occurs when the magnitude of the reactions to the relevant and the control questions are similar.

 $^{^2}$ The National Research Council did not report accuracy in terms of percentage correct decisions. This is because percentage correct decisions rely on an arbitrary cutoff point. The choice of where the cut-off is placed depends on the preference to reduce either the false positive ratio or the false negative ratio. Rather, the National Research Council expressed accuracy in terms of the area under the Receiver Operating characteristic Curve (ROC *a*).

³ This conclusion was highly similar to a U.S. government report that was published 20 years earlier (Office of Technology Assessment, 1983). This report concluded that "...the polygraph detects deception at a rate better than chance, but with error rates that could be considered significant."

guilty persons escape than that one innocent suffer (Blackstone, 1882; see also Volokh, 1997)."

Pseudoscience

The National Research Council report cited above also contains some other noteworthy conclusions. For example, the Council concluded that 'Research on the polygraph has not progressed over time in the manner of a typical scientific field. It has not accumulated knowledge or strengthened its scientific underpinnings in any significant manner. (p. 213)'. This absence of progress can, at least in part, be explained by the fact that the use of the polygraph is almost exclusively in the hands of practitioners who lack ties to academia. Historically, the CQT has been largely developed outside academia, and the examiners administering the tests are most often law enforcement officials without any academic background in psychology, physiology or psychophysiology. Without such a background, it is difficult to fully comprehend the scientific literature on polygraph testing. Hence, polygraph examiners almost exclusively rely on what they are trained at polygraph schools: That the CQT is highly accurate.⁴ When they start conducting tests in the field, selective feedback further strengthens their perception of the CQT being neatly infallible (Vrij, 2008). A closer look at how the CQT is employed in the field may help to understand this. Typically, a 'deception indicated' test outcome is followed by an interrogation during which the suspect is confronted with the test outcome. If the suspect confesses during this interrogation, this provides the examiner with the feedback that the test outcome was correct. If it was an innocent suspect who failed the test, a confession is unlikely. This does not, however, prove the test outcome wrong, as the examiner may

⁴ The British and European Polygraph Association, for example claims that 'research has shown that the accuracy of computerized polygraph testing is 98 %. See http://europeanpolygraph.org/faqs.htm

uphold the belief that the suspect is indeed guilty, but he or she did just not confess. If the outcome of the test is 'no deception indicated', no subsequent interrogation takes place. In case of a guilty suspect escaping detection, this error will go unnoticed, as the absence of an interrogation ensures the absence of a confession. Thus, this mechanism of selective feedback ensures occasional feedback on correct decisions while preventing feedback in the cases of an incorrect outcome, thereby explaining the perceived infallibility of the polygraph by examiners. It also makes the following conclusion from the NRC report a little less remarkable: 'What is remarkable, given the large body of relevant research, is that claims about the accuracy of the polygraph made today parallel those made throughout the history of the polygraph: practitioners have always claimed extremely high levels of accuracy, and these claims have rarely been reflected in empirical research.' (p. 107).

Sex offenders and job applicants

The polygraph test is also becoming increasingly popular in the treatment and monitoring of sex offenders. Polygraph tests are used to monitor offenders' activities in the community (e.g., during a parole), as well as to gain a more comprehensive understanding of their historical sexual interests and behaviours (English, Jones, Patrick, & Pasini-Hill, 2003). In contrast to tests in criminal investigations where the offense is known and the question is whether the suspect committed it, the sex offender is questioned about incidents of which it is unknown whether they have taken place at all. This type of test bears close resemblance to the application of the polygraph in personnel screening, used in the United States to screen job applicants and monitor employees of government agencies whose work involves security risks (e.g., FBI applicants or nuclear

scientists; Krapohl, 2002). The relevant questions are necessarily phrased in a very broad way (e.g., "Have you had unsupervised contact with children over the last 3 months?"). Relevant questions and control questions (e.g., "Have you done anything over the last 3 months that would concern your probation officer?"; Grubin et al., 2004 p. 213) become more similar, test outcomes more ambiguous and diagnostic decisions less accurate (National Research Council, 2003).

Despite these concerns, several reports have suggested that the polygraph is highly successful in obtaining previously undisclosed information (e.g., Ahlmeyer, Heil, McKee, & English, 2000; Emerick & Dutton, 1993; English, Jones, Patrick, & Pasini-Hill, 2003; Grubin, Madsen, Parsons, Sosnowski, & Warberg, 2004; Wilcox & Sosnowski, 2005). This includes information on the number of previous offences, age of the first offence, number of victims and prevalence of high risk behaviors. It is important to realize that the capability of the polygraph to obtain new information is more related to its intimidating effect than to its accuracy. According to Abrams and Abrams (1993), for example, there are three points in time when sex offenders can disclose information: (1) when they are told that they will face a test in the near future, (2) during the pre-test interview, and (3) during the confrontation with the test outcome. Note that the first two points are before the polygraph test is actually conducted. Indeed, Grubin and his colleagues (2004) found that most information was disclosed during the pre-test interview, well before the sensors of the polygraph were actually connected. Apparently, the expectation of an upcoming polygraph test is sufficient to make offenders disclose information. This means that disclosure of new information may have little to do with the polygraph as a method for the detection of deception per se. More likely, it is the questioning and the intimidation by the lie detector that makes the examinee bring up new information (see also Gannon, Keown, & Polaschek, 2007; Meijer, Verschuere, Merckelbach, & Crombez, 2008).

Little is known about the effects of newly disclosed information on treatment success. In a first study investigating this issue, McGrath and colleagues (2007) compared 5-year re-offense rates of a group of 104 adult male sex offenders who received community cognitive-behavioral treatment, correctional supervision, and periodic polygraph compliance exams with a matched group of 104 sex offenders who received the same type of treatment and supervision services but no polygraph exams. In line with earlier studies, polygraph testing resulted in the disclosure of previously withheld high risk behaviors, and the vast majority of the treatment providers and supervision officers rated the tests as "helpful" or "very helpful". Still, their data did not provide much support for an effect of polygraph tests on recidivism. The number of individuals in the polygraph group charged with committing a new non-sexual violent offense was significantly lower than in the no polygraph group (2.9% versus 11.5%). However, there were no significant between-group differences for the number of individuals charged with a new sexual (5.8% versus 6.7%), or other (non-sexual and nonviolent) offense (35.6% versus 29.8%). There was also no significant difference in overall recidivism (sexual, violent, or other; 39.4% versus 34.6%). Also, the number of individuals known to have violated their supervision conditions did not differ between groups (51.9% versus 45.2%), nor did the number of individuals who returned to prison (47.1% versus 38.5%). Based on these findings, the authors argued that the polygraph's "widespread use has far outpaced empirical examination of its effectiveness" (McGrath, Cumming, Hoke, & Bonn-Miller, 2007, p. 391).

Concealed Information

In the late fifties, professor of psychology and strong opponent of the CQT, David Lykken, developed an alternative application of the polygraph. This alternative application was, in fact, first described by Münsterberg (1908) and later named the Guilty Knowledge Test by David Lykken (1959). It is nowadays commonly referred to as the Concealed Information Test (CIT). During a CIT, test questions do not directly address the incident under investigation. Rather, all questions concern details of the crime, presumably only known to the police and to the perpetrator. Answer alternatives to these questions are presented serially, while physiological signals are recorded. These answer alternatives include the correct answer, but also several plausible but incorrect answers (e.g., "Was the victim killed with a ... (a) gun, (b) knife, (c) rope, (d) bat, (e) ice pick?"). For an innocent suspect, all alternatives are equally plausible,⁵ and will elicit similar physiological responses. For a guilty suspect, on the other hand, the correct alternative is salient, and will elicit an enhanced physiological response. When multiple questions, each pertaining to different details, are presented to the suspect, and he or she shows a pattern of stronger responding to the correct alternative, knowledge of intimate details of the crime is determined, from which involvement can be inferred.

One of the main advantages of the CIT is that the probability of a false positive test outcome is fully under the control of the examiner. A false positive test outcome

⁵ There are two ways to ensure that the different alternatives are equally plausible. First, like with an Oslo confrontation, all alternatives can be presented to a panel of naïve participants with the question to guess which alternative is correct (Doob & Kirschenbaum, 1973). Second, the alternatives can be presented to the suspect before the actual test in a previewing session (Verschuere & Crombez, 2008).

means that, merely by chance, a pattern of stronger responding to the correct alternatives has occurred. The probability of this happening depends on two factors. The first factor concerns the test's properties; the false positive probability is inversely related to the number of questions and the number of answer alternatives per question. The second factor that determines false positive probability is how one defines 'a pattern of stronger responding'. When a guilty test outcome requires the suspect to respond maximally to the correct alternatives of all five questions, the probability of this happening by chance, is smaller than when one requires a maximal response for only three out of the five correct alternatives. This control over false positive probability has important implications. For one thing, it allows the examiner to set the false positive probability at an arbitrary low level, as prescribed by legal doctrine in most civilized countries. It also allows for calculation of the probability that a guilty test outcome is incorrect. This is essential information if an incriminating test outcome is introduced in court proceedings, as it allows for proper weighing.

The CIT originally described by Lykken (1959) used only skin conductance responding (SCR) as the dependent measure. This measure has by far received the most attention in CIT research, and has been shown robust in discriminating between guilty and innocent participants (see Table 1). Elaad (1998), for example, reviewed 15 mock crime studies and found average detection rates of 81% for guilty examinees and 96% for the innocent. Similar accuracy rates were reported by Ben-Shakhar and Furedy (1990), who reviewed 10 mock crime studies and found detection rates of 84% and 94% respectively. A more recent review showed similar results with successful detection of 76% of participants with concealed knowledge and slightly lower detection, 83%, of those without (MacLaren, 2001). The two available field studies (Elaad, 1990; Elaad, Ginton, & Jungman, 1992) show equally high detection of innocent suspects (98% and 94%, respectively), but somewhat lower detection accuracy among guilty suspects (42% and 76%, respectively). The latter was, at least in part, attributed to the suboptimal format of the tests, as they contained only a low number of questions. The most extensive meta-analysis so far was performed by Ben-Shakhar and Elaad (2003). Eighty studies were included, yielding an average effect size (*d*) of 1.55, which is typically considered a very large effect size (Cohen, 1988). Furthermore, an area under the Receiver Operating Characteristic curve were even higher under conditions that best mimicked real life application (3.12 and .95, respectively). Finally, in its 2003 report, the National Research Council selected 13 studies, yielding an area under the Receiver Operating Characteristic curve of .88.

Based on the controllable false positive rate and its firm grounding in orienting theory (Verschuere, Crombez, De Clercq, & Koster, 2004), many researchers feel that the CIT is ready for field application (Ben Shakhar, Bar Hillel, & Kremnitzer, 2002). Still, it is only rarely used, with Japan being the only exception (Hira & Furumitsu, 2002; Nakayama, 2002). This limited use can, at least in part, be attributed to the difficulty of formulating sufficient appropriate test items. These items need to fulfill two requirements. To begin with, the details asked need to be known to the culprit and the investigating authorities. Second, these items must not be known to an innocent suspect, e.g., through the media. Estimations of the number of cases that meet these criteria are given by Podlesney (1993; , 2003). This author examined the files of FBI examinations in

which a COT was performed, for usable CIT items, and found that this was the case in 13% and 7%, respectively. At first glance, these statistics do not look very encouraging. It should be noted, however, that Podlesney's estimation is based on a retrospective review of case records. This way of post-hoc determination of possible test is likely to underestimate potential application, because the information in these records was not selected with a CIT in mind (see also Lykken, 1998). Furthermore, discussions about applicability seem to be dictated by the issue of whether the CIT can fully replace the CQT (e.g., Ben-Shakhar, 1991). This, however, should not be the primary question. The relevant question that should be answered is whether a CIT can be applied in a substantial proportion of the cases. With this question in mind, Podlesney's data sketch a much more optimistic picture. After all, it shows that in difficult cases (otherwise a CQT would not have been necessary) a CIT would have been possible in approximately 10% of the cases. This figure should be regarded as the lower bound estimate as it is based on incomplete file records. In a regular forensic context, the applicability of the CIT seems amply enough to yield a positive cost / benefit analysis. The large scale use of the CIT in Japan has, in any case, shown that successful implementation is indeed possible (Hira & Furumitsu, 2002; Nakayama, 2002).

The future

The search for a more accurate lie detection procedure continues. Many think that with the advancement of technology, polygraph tests will become more accurate. The review by the National Research Council, however, contradicts this idea. In their review, plotted the accuracy of the selected polygraph studies was plotted against the year of publication (National Research Council, 2003, p. 346). If technological advancement would lead to

an increased accuracy, a positive trend should be apparent. This was not the case. Erroneous outcomes of the CQT occur because the polygraph traces do not allow for distinguishing between an innocent's fear of false detection and a guilty's fear of detection. This problem is not solved by technologically more sophisticated measuring device or more advanced scoring algorithms. As long as the question format of the CQT remains the same, polygraph test will not reach extremely high accuracy (National Research Council, 2003).

With regard to the CIT, problems are only of a practical nature. It is unlikely that investigative authorities will successfully implement the method on their own, and researchers should be willing to invest in assisting the investigative authorities with setting up an infrastructure that allows for conducting CITs. Because of the low probability of false positive errors, the CIT can be tested under field situations relatively safely. After all, the risk of incriminating an innocent suspect is small.

Concluding Remarks

The use of the CQT for the detection of deception is highly controversial. Strong claims about its accuracy and utility almost exclusively comes from practitioners. Academics, on the other hand, are mostly skeptical. Practitioners who are involved in the use of polygraph tests, whether it is in law enforcement, sex offender treatment or any other application should keep the following in mind that the inner logic of the CQT is weak, and an innocent examinee may also respond to the relevant questions. In specific incident (i.e., crime) testing, where the offense is known, it's accuracy has been shown to exceed chance level. The error rate is, however, substantial, and in screening situations (i.e., sex offender monitoring and treatment), where the offense is unknown, the error rate is most likely even higher. The problems associated with the CQT are of a logical nature, and will not be solved by new sensors or advanced scoring algorithms. Due to its low false positive rate, the CIT can safely be applied in the field. For successful application, researchers need to assist investigative authorities.

References

- Ahlmeyer, S., Heil, P., McKee, B., & English, K. (2000). The impact of polygraphy on admissions of victims and offenses in adult sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 12(2), 123-138.
- Ben-Shakhar, G. (1991). Future prospects of psychophysiological detection: Replacing the CQT by the GKT. In J. R. Jennings, P. K. Ackles & M. G. H. Coles (Eds.), *Advances in Psychophysiology* (Vol. 4, pp. 193-199). London: Jessica Kingsley.
- Ben-Shakhar, G. (2008). The case against the use of polygraph examinations to monitor post-conviction sex offenders. *Legal and Criminological Psychology*, 13, 191-207.
- 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(5), 527-541.
- Ben Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, 88(1), 131-151.
- Blackstone, W. (1882). Commentaries on the laws of England (3rd ed.). London: Murray.
- Cohen, J. E. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Doob, A. N., & Kirschenbaum, H. M. (1973). Bias in police lineups-partial remembering. *Journal of Police Science and Administration*, 1, 187-293.
- Elaad, E. (1990). Detection of guilty knowledge in real-life criminal investigations. *Journal of Applied Psychology*, 75(5), 521-529.
- Elaad, E. (1998). The challenge of the Concealed Knowledge Polygraph Test. *Expert Evidence*, *6*(3), 161-187.
- Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real-life criminal guilty knowledge tests. *Journal of Applied Psychology*, 77(5), 757-767.
- Emerick, R. L., & Dutton, W. A. (1993). The effect of polygraphy on the self report of adolescent sex offender: Implications for risk assessment *Annals of Sex Research*, 6, 83-103.
- English, K., Jones, L., Patrick, D., & Pasini-Hill, D. (2003). Sexual offender containment: Use of the postconviction polygraph. *Annuals New York Academy of Sciences*, 989, 411-427.
- Fiedler, K., Schmid, J., & Stahl, T. (2002). What is the current truth about polygraph lie detection. *Basic and Applied Social Psychology*, 24(4), 313-324.
- Gannon, T. A., Keown, K., & Polaschek, D. L. (2007). Increasing honest responding on cognitive distortions in child molesters: The bogus pipeline revisited. Sex Abuse, 19, 5-22.
- 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(3), 209-222.
- Hira, S., & Furumitsu, I. (2002). Polygraphic examinations in Japan: Application of the guilty knowledge test in forensic investigations. *International Journal of Police Science & Management*, 4, 16-27.

- Honts, C. R. (2004). The psychophysiological detection of deception. In P. A. Granhag & L. A. Strömwall (Eds.), *The detection of deception in forensic contexts* (pp. 103-126). Cambridge: Cambridge University Press.
- Iacono, W. G. (2008). Effective policing Understanding how polygraph tests work and are used. *Criminal Justice and Behavior*, *35*, 1295-1308.
- 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.
- Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, 43, 385-388.
- Lykken, D. T. (1998). A tremor in the blood. New York: Plenum Press.
- MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *Journal of Applied Psychology*, 86(4), 674-683.
- McGrath, R. J., Cumming, G. F., Hoke, S. E., & Bonn-Miller, M. O. (2007). Outcomes in a community sex offender treatment program: A comparison between polygraphed and matched non-polygraphed offenders. *Sex Abuse*, *19*(4), 381-393.
- Meijer, E. H., & van Koppen, P. J. (2008). Lie detectors and the law: The use of the polygraph in Europe. In D. Canter & R. Zukauskiene (Eds.), *Psychology and law: Bridging the gap* (pp. 31-50). Aldershot, UK: Ashgate Publishing.
- Meijer, E. H., Verschuere, B., Merckelbach, H. L., & Crombez, G. (2008). Sex offender management using the polygraph: A critical review. *International Journal of Law* and Psychiatry, 31(5), 423-429.
- Munsterberg, H. (1908). On the witness stand. New York: The McClure Company.
- 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. Division of behavioral and social sciences and education.* Washington, DC: The National Academic Press.
- Offe, H., & Offe, S. (2007). The comparison question test: Does it work and if so how? *Law and Human Behavior*, *31*(3), 291-303.
- Office of Technology Assessment. (1983). *Scientific validity of polygraph testing: A research review and evaluation*. Washington, DC: US Government Printing Office.
- Podlesney, J. A. (1993). Is the guilty knowledge technique applicable in criminal investigations? A review of FBI case records. *Crime Laboratory Digest*, 20, 57-61.
- Podlesney, J. A. (2003). A paucity of operable case facts restricts applicability of the guilty knowledge technique in FBI criminal polygraph examinations. *Forensic Science Communications*, 5(3), Retrieved December, 2, 2008 from <u>http://www.fbi.gov/hq/lab/fsc/backissu/july2003/podlesny.htm</u>.
- 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.
- Reid, J. E. (1947). A revised questioning technique in liedetection tests. *Journal of Criminal Law and Criminology*, *37*, 542-547.

- Verschuere, B., & Crombez, G. (2008). Deja vu! The effect of previewing test items on the validity of the Concealed Information polygraph Test. *Psychology, Crime & Law, 14*, 287-297.
- Verschuere, B., Crombez, G., De Clercq, A., & Koster, E. H. (2004). Autonomic and behavioral responding to concealed information: Differentiating orienting and defensive responses. *Psychophysiology*, 41(3), 461-466.

Volokh, A. (1997). n Guilty men. University of Pennsylvania Law Review, 146, 173-211.

Vrij, A. (2008). Detecting lies and deceit. Pitfalls and opportunities. Chichester: Wiley.

Wilcox, D. T., & Sosnowski, D. E. (2005). Polygraph examination of British sexual offenders: A pilot study on sexual history disclosure testing. *Journal of Sexual Aggression*, 11(1), 3-23.

Figure 1. Overview of the accuracy figures yielded by laboratory and field studies for the Control Question and Concealed Information polygraph test.

		Guilty participant / suspect			Innocent participant / suspect			
	n	Test outcome	Test outcome	Test outcome	Test outcome	Test outcome	Test outcome	
	studies	correct	incorrect	inconclusive	correct	incorrect	inconclusive	
		Control Question Test						
Laboratory studies								
Office of Technology Assessment (1985)	12	74	7	19	60	16	24	
Kircher et al. (1988)	14	74	8	18	66	12	22	
Ben-Shakhar and Furedy (1990)	9	80	7	13	63	15	22	
Honts (2004)	11	82	7	11	83	10	7	
Field studies								
Office of Technology Assessment (1985	10	87	11	2	75	19	6	
Ben-Shakhar and Furedy (1990)	9	84	13	3	72	23	5	
Honts (2004)	4	89	1	10	59	12	29	
		Concealed Information Test						
Laboratory studies								
Ben-Shakhar and Furedy (1990)	10	84	16		94	6		
Elaad (1998)	15	81	19		96	4		
Lykken (1998)	8	88	12		97	3		
MacLaren (2001)	22	76	24		83	17		
Field studies								
Elaad (1990)	1	42	58		98	2		
Elaad, Ginton and Jungman (1992)	1	76	24		94	6		

Box 1. An example of the pre test interview as given by Offe and Offe (2007) in an experiment where the theft of a 0 voucher is the incident under investigation.

I am going to ask you some questions to find out what your history concerning this matter looks like. I want to give you the reason, too, why I have to ask you such indiscrete questions. I want to find out whether one would consider you capable of an action such as removing a voucher for $50 \in$ based on your history or not. In a nutshell, I want to know whether such an action, taking something of monetary value out of a closed room fits your personality profile or not.

These personal questions also have to be answered entirely truthfully. The more of these questions you can truthfully negate, the better it is for you, because then one can say that such an action does not fit your personality profile. If, however, you have to truthfully answer yes, then I will have to continue asking what the context was, so that I can get an impression of whether these were small and harmless delinquencies or whether there were some serious ones as well. Depending on what you tell me, it may begin to become imaginable that you may have done what we are talking about here as well.

The personal questions have nothing to do with whether you have taken the voucher or not. In order to make that clear from the beginning of every question, each personal question will start with: "In the first 25 years of your life, …," so that you will know right away, "this is about my past." For the result of the polygraph examination it is important, that you answer these questions truthfully as well.