UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Memory Detection: Past, Present and Future

Geven, L.M.; Ben-Shakhar, G.; Kindt, M.; Verschuere, B.

DOI
10.1007/978-3-319-96334-1_19

Publication date
2019

Document Version
Submitted manuscript

Published in
The Palgrave Handbook of Deceptive Communication

Link to publication

Citation for published version (APA):
Geven, L. M., Ben-Shakhar, G., Kindt, M., & Verschuere, B. (2019). Memory Detection: Past, Present and Future. In T. Docan-Morgan (Ed.), *The Palgrave Handbook of Deceptive Communication* (pp. 367-383). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-96334-1_19

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:09 Mar 2023

Memory Detection: Past, Present, and Future

Linda Geven[a,b], Gershon Ben-Shakhar[b], Merel Kindt[a] and Bruno Verschuere[a]

[a]Department of Clinical Psychology, University of Amsterdam, The Netherlands

[b]Department of Psychology, Hebrew University of Jerusalem, Israel

Geven, L. M., Ben-Shakhar, G., Kindt, M., & Verschuere, B. (2018). Memory detection: Past, present, and future. In T. Docan-Morgan (Ed.), *The handbook of deceptive communication*. London, UK: Palgrave Macmillan.

Abstract

The Concealed Information Test (CIT) aims to detect the recognition of concealed knowledge in an interviewee by presenting a series of multiple-choice questions while measuring several psychophysiological (e.g., skin conductance) or behavioral (i.e., reaction times) responses. When a suspect consistently shows distinct responses to the critical (e.g., crime-related) items compared to the neutral control items, knowledge is inferred. This chapter provides an overview of memory detection using various response measures, including research findings and the underlying mechanisms. Although available data confirm the validity of the CIT, there is quite a gap between these laboratory studies and realistic criminal investigations. Possible ways to tackle challenges that lie ahead, including field validity, leakage of critical information to innocent suspects, and discovering intentions are discussed.

Key words: Concealed Information Test (CIT), Memory detection, Deception, External Validity, Psychophysiology

Memory Detection: Past, Present, and Future

It was November 1985 when the bodies of two sexually assaulted, murdered and mutilated children, a seven-year-old girl and her eight-year-old brother, were found in Plainfield, New Jersey. Byron Halsey, the boyfriend of the children's mother, quickly became the main suspect in the high–profile investigation. Was Mr. Halsey the true perpetrator of this crime? And could a lie detector help the case?

A Short History On Lie Detection

Most people are familiar with the Pinocchio effect; the nose of the wooden puppet instantaneously and observably grew whenever he told a lie. This famous story does not stand alone in the history of detecting deception. In ancient India, suspects were asked to chew raw rice and those who could not spit out the rice were in big trouble. It was believed that liars have a dry mouth, therefore making the rice stick to the tongue. Similar methods were used by the Bedouins of Arabia, using an even more brutal method. If the tongue of the suspect would stick to a burning hot iron, deception was indicated (Lykken, 1998; Trovillo, 1939).

Whether it is a growing nose or a dry mouth, no single bodily response has been established to be uniquely related to lying (Vrij, 2008). As a result, many Indians and Bedouins might have been wrongfully convicted. The problem lies in the flawed theory underlying the idea to use a dry mouth as a cue to deception. What seems to be forgotten is the fact that even an innocent person fearing to be falsely accused could experience increased stress and enhanced bodily responses when facing a lie detection test. Some contemporary deception detection methods, such as the Control Question Polygraph Test (CQT; Reid, 1945), are still based on the idea that fear or stress responses reveal deception. Since an incorrect outcome may put an entire investigation on the wrong track (Kassin, Bogart, & Kerner, 2012), invalid lie detection tests based on stress–induced cues should be avoided.

Byron Halsey, suspected of the molestation and brutal murder of the two children, was convicted to two life sentences plus 20 years, after failing the polygraph examination. Importantly, on key crime details such as the location of the bodies and the modus operandi, Halsey initially gave an incorrect narrative before guessing the correct manner of death and confessing to the crimes in a tense interrogative setting. More than two decades later, after spending most of his prison time in solitary confinement for his own safety, post-conviction DNA testing not only proved Halsey's innocence, but also implicated the true perpetrator of the horrific crime.

According to William Blackstone in his book *Commentaries on the Laws of England* (1830), exonerating ten guilty individuals is deemed better than wrongfully incarcerating a single innocent person. Since then, preventing false positive errors (i.e., mistakenly identifying an innocent person to be guilty) forms the basis of our legal system in which the scales of justice are tilted in favor of the accused unless sufficiently proven guilty. Deception detection methods should therefore ideally not only reach good sensitivity (i.e., proportion of correctly detected guilty suspects based on the test outcome), but particularly high levels of specificity (i.e., proportion of correctly identifying the innocents), in order that errors as the one made in the case of Byron Halsey are avoided. In high stake situations, such as in the case of criminal proceedings, accurate and reliable deception detection techniques are an absolute necessity. Methods to detect deceit should be based on a sound scientific framework as to reliably indicate possible involvement in a crime and avoid wrongful incarceration.

Memory Detection

The fundamental difference between flawed polygraph tests fixated on deception-induced stress (Ben-Shakhar, 1991; Ben-Shakhar, Bar-Hillel, & Lielblich, 1986; Lykken, 1991) and methods designed to detect memory traces, is that the latter methods focus on detecting recognition of intimate crime details rather than deception. While lie detection attempts to determine deception by interpreting answers to interrogational questions such as “Did you kill the two children?”, the purpose of the Concealed Information Test (CIT; first introduced as the Guilty Knowledge Test by Lykken, 1959, 1960) is to verify whether the suspect is aware of certain crime-related information, for instance whether the murder weapon was a bomb, a firearm, or a knife. This method is therefore labeled a memory detection test rather than a lie detection test.

The objective of the CIT is to verify whether the suspect possesses crime-related information that only the perpetrator would be aware of. The method requires that the examiner determines a number of established facts from the investigation which only the true culprit will be able to recognize. Then, the examiner creates a CIT resembling a multiple-choice test with several questions, such that each question is comprised of one detail of the crime in question, and several neutral control items. In the case of a homicide, for instance, the CIT might involve questions concerning the murder weapon and the location of the victim in the crime scene. For example ‘How was the victim murdered? a) by beating, b) by stabbing, c) by drowning, d) by shooting or e) by poisoning’ or ‘Where was the victim attacked? a) bathroom, b) kitchen, c) bedroom, d) garden, e) living room’. For each question, there is only one item that reflects the correct feature of the crime under investigation (i.e., the critical or crime-related item, called the probe). The other options are neutral control items from the same category as the relevant item (i.e., called the irrelevants). These irrelevant items are chosen carefully, such that all options would seem equally plausible to unknowledgeable individuals. As a result, to innocent suspects who are unaware of the crime's details, all items will trigger similar responses. On the other hand, guilty suspects tend to react significantly different to the recognized crime-related detail than to the irrelevant items. This differential response (i.e., probe minus irrelevant response, labeled the CIT effect) indicates critical knowledge of the crime in question, which should lead to further investigation of the suspect (Lykken, 1974, 1998).

Imagine the case of the double child murder in New Jersey. If Byron Halsey would have been guilty, it is likely that he remembered the location of the victims' bodies. Therefore, he would have shown a differential response to the correct alternative (i.e., basement). On the other hand, since the suspect was in fact innocent and therefore did not know and could not infer which alternative reflected the true feature of the crime, he would show similar responses to all items. By using several CIT questions, each with 4 or 5 alternatives and several dependent measures, the probability of a false positive outcome could be controlled by the investigators. The main advantage of the CIT over methods focusing on deception is its use of proper controls. The CIT establishes a within-person control in which responses to the critical alternative are compared to an estimate of the response to the correct alternative if the person would be in fact innocent (i.e., the irrelevant options).

Moreover, stress-induced mental states driven by the potential consequences of failing the test are expected to influence both relevant and irrelevant alternatives similarly. Thus, whether the suspect is calm, aroused or even frightened, it is still expected that the response to the critical crime-related item is stronger than to the alternatives whenever the suspect recognizes the correct answer. Likewise, if the suspect is in fact innocent and unaware of the crime-related items, neither the present emotional state nor the possible consequences of failing the test can influence the CIT outcome, since it affects responses to both critical and control items. In addition, the CIT is a scientific approach to deception detection, substantiated by extensive theoretical and applied research (e.g., Ben-Shakhar, 2012; Verschuere, Ben-Shakhar, & Meijer, 2011).

Underlying Theory

The Orienting Response (OR) has long been the dominant hypothesis for the CIT effect (see Ben-Shakhar, 1977; Lieblich, Kugelmass, & Ben-Shakhar, 1970; Lykken, 1974). A combination of a physiological and behavioral responses in reaction to an external novel stimulus, forms the basis of the OR. Already in 1927, while working on his famous classical conditioning experiments, Pavlov noted that the dogs in his laboratory shifted their attention to new visitors. But it did not take long before their interest in this novel person would decline, resulting in a decrease of the dogs' investigatory response (Sokolov, 1963). This incidental finding demonstrated that the initial orienting response may gradually habituate in magnitude. However, when the stimulus holds a special significance to the subject, an enhanced OR can be observed (Sokolov, 1963). Significant stimuli have also been proven to be more resistant to habituation. Changes in the magnitude of the OR therefore allow for differentiation between salient and neutral stimuli (Bradley, 2009; Gamer, 2011; Lykken, 1974). This effect also forms the basis of the Concealed Information Test. The critical crime detail (e.g., strangulation of the victim) holds a very significant meaning to the guilty individual, but not to the innocent suspect. Therefore, a consistently stronger response to 'strangulation' compared to the control items (e.g., 'drowning', 'poisoning', etc.), is an indication that the suspect has knowledge about the crime in question and should be further investigated.

More recently, response inhibition was found to also underlie the CIT effect (klein Selle, Verschuere, Kindt, Meijer, & Ben-Shakhar, 2016, 2017a; Suchotzki, Verschuere, Peth, Crombez, & Gamer, 2015; Verschuere, Crombez, Koster, Van Bockstaele, & De Clercq, 2007). Since it is assumed that the truthful answer is the natural and default response to a question, lying requires a significant amount of cognitive resources in order to actively suppress the truth (Suchotzki et al., 2015). Response inhibition is thus required to prevent the truth from being exposed in the CIT and to give a deceptive response instead. For the CIT rationale, it is reasoned that in order to remain undetected, a guilty suspect must suppress the increased arousal associated with recognition of the critical item. However, this effort has the paradoxal consequence (Pennebaker & Chew, 1985) that it further increases physiological responses to the critical items.

Response Measures

Recognition of critical items in the CIT can be assessed by autonomic and behavioral measures, as well as brain–related measures, such as functional magnetic resonance imaging (fMRI) or electroencephalography (EEG). For all measures, the key factor in memory detection is the differential response to the critical items compared to the irrelevant options as an indicator of recognition. Autonomic nervous system (ANS) measures such as skin conductance, heart rate and respiration have been used since the beginning of memory detection. While measuring amplitudes of the galvanic skin response elicited by items in the CIT, larger skin conductance responses (SCR) upon probe presentation were found for individuals attempting to conceal information (Lykken, 1960). Moreover, respiration (RLL), measured with pneumatic straps around the chest and abdomen, is smaller upon recognition of the relevant items compared to neutral items. Similarly, cardiovascular measures can indicate concealed information. Phasic heart rate (HR), measured with electrodes on the chest or with infrared at the fingertip, decreases within 15 seconds after presentation of the critical item, compared to irrelevant alternatives.

Does it Work?

Since the early 1960s, there is ample evidence for successfully detecting crime–related knowledge and discriminating between guilty and innocent individuals with the CIT. An early laboratory study on the validity of the CIT was conducted in 1959 by David Lykken, laying a promising foundation for future research. In an attempt to mimic real–life situations in which memory detection tests could be meaningful, a mock-crime paradigm was used. By measuring and ranking the amplitude of the galvanic skin response upon presentation of the probe and irrelevant alternatives for both guilty and innocent participants, around 90% of the participants were classified correctly. In a meta-analysis (Ben-Shakhar & Elaad, 2003), the validity of 42 subsequently executed mock-crime experiments revealed a very large average effect size (Cohen´s *d*) of 2.09 (0.80 is considered a large effect size by Cohen, 1988). These results confirmed that the SCR measure can accurately detect relevant information and differentiate between individuals with and without knowledge of the critical mock-crime details.

In addition to skin conductance, various other experiments added different measures of the autonomic nervous system such as changes in respiration and heart rate. In most observations, the SCR outperformed other measures in detecting recognition (see for example Podlesny & Raskin, 1978; Balloun & Holmes, 1979). While the SCR remains the most valid single autonomic measure, an accumulation of all three was found to provide incremental evidence for the effectiveness of the CIT (Gamer, Verschuere, Crombez, & Vossel, 2008). More recently, meta-analytic results reconfirmed the validity of these psychophysiological measures to detect the presence or absence of crime-related knowledge in a suspects' memory (Meijer, klein Selle, Elber, & Ben-Shakhar, 2014). Again, large effect sizes were found for the SCR, RLL and HR (1.55, 1.11, and 0.89, respectively).

Another shift of interest took place in the last two decades towards the potential of the response latency measure for detecting deception and concealed knowledge. Relying on reaction times (RTs) to distinguish between innocents and individuals withholding critical information requires a slightly different procedure, yet the response difference between the probe and several irrelevant options remains essential. Initial research using handheld stopwatches did not find strong effects for reaction times as an index of deception (see Luria, 1932; Marston, 1920), but computerized methods have led to a renewed research interest in the use of response latency to detect deception.

While measuring RTs, participants in a laboratory study (Seymour, Seifert, Shafto, & Mosmann, 2000) had to indicate whether they recognized the stimuli presented in the CIT by pressing one of two response keys. Critical details from the committed mock-crime were intermixed with neutral items. Upon measuring response latencies for denying knowledge of the probe in comparison to irrelevant words, 23 out of 27 participants were correctly classified as guilty. A recent meta-analysis based on studies relying on computerized RT measures showed the potential of the RT–CIT (Cohen's *d* of 1.30; Suchotzki, Verschuere, Van Bockstaele, Ben-Shakhar, & Crombez, 2017).

Interestingly, recent insights suggest that different response measures may be driven by different mechanisms. Specifically, it has been suggested that some measures (e.g., elevated skin conductance) may be mostly related to the concealed item drawing attention (i.e., orienting response), whereas other measures (e.g., the drop in heart rate, more shallow breathing cycles, and slowing of reaction time) reveal the subsequent deliberate concealment of the recognition (i.e., response inhibition; klein Selle et al., 2016, 2017a; Rosenfeld, Oszan, & Ward, 2017; Suchotzki et al., 2015).

Whereas laboratory research allows for a controlled environment and manipulation of isolated variables, questions can be asked about the generalizability of the results to real-life cases (i.e., ecological validity). Establishing accuracy rates for memory detection in a realistic situation such as criminal investigations, requires validity studies conducted in authentic settings. Therefore, the few field studies that were reported may provide an additional insight regarding the external validity of laboratory experiments designed to assess the validity of the CIT. These studies conducted with real suspects, for whom the stakes are high and who are motivated to avoid detection, produced mixed results. Specifically, Elaad (1990) found a very high accuracy rate (98%) for discerning innocent suspects with the SCR measure, but the rate for correctly identifying guilty suspects was much lower than expected (42%). In a second field study (Elaad, Ginton, & Jungman, 1992), measuring RLL in addition to SCR, the results were a bit more promising, although still far from optimal. Both measures separately could detect innocent and guilty suspects with 97% and 53% accuracy, respectively, yet combining SCR with RLL led to increased detection accuracy (76%) for guilty suspects. However, it is worth noting that the CITs applied in these field studies were not optimal. First, the authors used a scoring procedure that is nowadays replaced by improved computational systems in which multiple measures can be combined and standardized. Second, the number of questions was much smaller than recommended and third, the CITs were administered immediately after a CQT, which might have attenuated the sensitivity of the measures due to habituation effects. Moreover,, when conducting field studies there might be other difficulties, mostly concerning the establishment of ground truth. In actual cases, it is very difficult to establish proof of whether the test outcome (either guilty or innocent) was in fact correct. Therefore, most field research uses confessions as the principle proof of actual guilt. Since this criterion is vulnerable to sampling biases (Iacono, 1991; Patrick & Iacono, 1991), the data should be considered with caution.

External Validity of CIT Studies

Due to the difficulties of establishing a solid ground truth criterion in realistic settings, several researchers have adopted an alternative approach for evaluating the external validity of CIT studies conducted in artificial laboratory conditions (for a review see, Ben-Shakhar & Nahari, 2018). Specifically, researchers have identified several factors that differentiate between laboratory and realistic environments and manipulated each factor in controlled experiments. In the following sections, we briefly review this research focusing on each of the identified factors: levels of arousal, motivation to avoid detection, and the influence of a delay between crime and CIT, as well as effects of real–life deception on external validity.

Level of Arousal

Clearly, suspects undergoing criminal investigations are much more aroused than subjects participating in laboratory experiments. Indeed, the average heart rate of examinees in real-life tests is much higher than that of laboratory examinees (Verschuere, Meijer, & De Clercq, 2011). Early CIT studies relied on the card test paradigm, where subjects pick a card from a deck, hide this information and a subsequent CIT is conducted to detect the hidden information. Evidently, this is very different from a realistic scenario in the legal field. More recent studies adopted the mock-crime paradigm, which seems to better approximate realistic crimes. In their meta-analysis, Ben-Shakhar and Elaad (2003) compared the CIT effect with the SCR measure obtained from card test and mock-crime experiments. They found a much larger CIT effect for mock-crime than for card test studies (2.09 vs. 1.35, respectively). However, the level of arousal experienced by mock-crime participants is still quite moderate and far below what real suspects may experience during realistic polygraph tests. To better tackle this question, Verschuere et al. (2011) conducted a card test study with suspects undergoing realistic police polygraph interrogations. The enhanced arousal level in this condition was confirmed by a higher baseline heart rate than typically measured with research participants in the laboratory. Even under these higher levels of arousal, the CIT-effect was found to exist in the field: HR, RLL, and SCR significantly changed upon presentation of the picked card as opposed to irrelevant cards. In a direct comparison of the CIT effect obtained in card-tests conducted during a realistic polygraph investigation and laboratory experiment, Zaitsu (2016) reported similar effects in both settings. Additional studies attempted to examine whether the level of arousal affects the outcomes of the CIT in controlled experiments. These studies, which employed different types of arousal manipulation, reported similar CIT effects in the high and low arousal conditions (Bradley & Janisse, 1981; klein Selle et al., 2017b; Kugelmass & Lieblich, 1966; Peth, Vossel, & Gamer, 2012).

Motivation to Avoid Detection

Another difference between actual examinations and laboratory simulations might be the motivation to avoid detection and appear innocent. While a guilty suspect might have sufficient reasons to keep up appearances, research participants obviously do not face comparable detrimental consequences. Motivational manipulations are commonly achieved by instructions (e.g., Gustafson & Orne, 1963), incentivizing participants for beating the polygraph (Bradley & Warfield, 1984), or punishing participants for an undesirable outcome (Lykken, 1974). Although experiments did not always reveal consistent findings (Gustafson & Orne, 1963, 1965; Horvath, 1978, 1979; Lieblich, Naftali, Shmueli, & Kugelmass, 1974), meta-analytic results support the notion that the SCR effect size increases when the motivation to avoid detection is high (Ben-Shakhar & Elaad, 2003; Meijer et al., 2014). Guilty suspects increasing their effort to deceive the test might therefore -paradoxically- show enhanced responses to the probe amongst irrelevant items and thus aid their own detection (e.g., motivational impairment hypothesis; DePaulo & Kirkendol, 1989). In contrast, CIT studies using response latency as the dependent variable do not seem to benefit from additional motivational instructions. In a meta-analytic study, liars under motivation instructions to appear innocent were detected equally adequate as a control condition (Suchotzki et al., 2017).

Delay Between Crime and CIT

While in the typical CIT experiment the test is administered immediately after participants were exposed to the critical items, realistic polygraph tests are often administered several weeks, or even months after the crime (Ben-Shakhar & Furedy, 1990). Naturally, memory declines with time and since the CIT is a memory detection test, this might pose as a pitfall in criminal investigations. Indeed, research findings confirm the weakening effect of a time delay on detection efficiency of the CIT (Carmel, Dayan, Naveh, Raveh, & Ben-Shakhar, 2003; Gamer, Kosiol, & Vossel, 2010; Nahari & Ben-Shakhar, 2011; Peth et al., 2012). However, this effect is mediated by the type of items used in the test. Whereas the memory for less important, peripheral items decays quite rapidly, detection of central items or the gist of an event (i.e., items that are directly associated with the crime, such as the weapon that was used) typically persist. In a real-life scenario, more reliable responses can be expected to a question about the murder weapon than to a question regarding the clothes of the victim. Examiners are therefore advised to use central crime details that are likely to be better encoded and more easily recalled.

The Free Choice to Commit a Crime, Deceive, and Conceal Information

Another important distinction between laboratory experiments and realistic criminal investigations is the deliberate aspect of criminal actions. Deception is commonly defined as a voluntary act (see Vrij, 2004), in which intention is a key factor. Yet, in laboratory studies on detecting deception, participants are often explicitly instructed to commit a staged crime and subsequently conceal knowledge (e.g., Lykken, 1959; Nahari & Ben–Shakhar, 2011). More recently, researchers have begun to explore the role of instructed versus spontaneous cheating and lying (Blakemore, Winston, & Frith, 2004; Kozel et al., 2005; Mohamed et al., 2006; Sip, Roepstorff, McGregor, & Frith, 2008). For instance, Nahari, Breska, Elber, klein Selle, and Ben-Shakhar (2017) gave participants a free 'choice' to decide whether to enact a mock crime or an innocent computerized task and compared those who choose to commit the mock crime with participants who were instructed to do so. The study revealed a similar CIT detection efficiency, based on SCR, RLL, and RT measures, in these two conditions.

However, deception, in all its complexity, can only be fully investigated when the decision to deceive is based entirely on the participants' own initiative. In an externally more valid paradigm, participants engaged in a trivia quiz and were provided with a monetary incentive for high accuracy performance. Participants were randomly allocated to either a condition where they were instructed to cheat on the quiz (mimicking the typical laboratory set–up) or to a condition in which they were provided with the opportunity to cheat using Google, yet without explicit instructions to do so. Assessments of their reaction times (Geven, Ben-Shakhar, Kindt, & Verschuere, 2018) and physiological responses (Geven, klein Selle, Ben-Shakhar, Kindt, & Verschuere, 2018) in the CIT revealed that both instructed and self–initiated cheaters showed a similar pattern upon recognition. The results indicate that the cognitive signature of lying is not restricted to explicitly instructed deception, but can also be observed for its real-life equivalent. These findings are highly encouraging from an ecological validity perspective, suggesting that when it comes to free choice and voluntary deception, the results of laboratory studies are a realistic reflection of the field.

Limitations of the CIT

In the previous sections, we emphasized the strength of the CIT as a scientific approach to memory detection, based on proper control questions with validity estimates generated from extensive laboratory research. We have also argued, when systematically examining several factors differentiating between the laboratory and realistic settings, that results of laboratory studies can be generalized. However, the CIT is by no means a perfect method and it is important to discuss its limitations. There are two main factors that might limit the validity of the CIT in realistic applications: The potential effects of countermeasures and the danger that critical crime–related items may contaminate innocent suspects.

Countermeasures

While increased motivation might enhance the CIT effect in the field, fear of detection can also tempt guilty suspects to use countermeasures (i.e., consciously alter physiological reactions during the CIT to avoid detection). Two strategies can be applied to diminish the expected probe-irrelevant difference during the polygraph examination. Suspects can either try to suppress their responses to the relevant crime details or artificially enhance responses to neutral, irrelevant items (Ben-Shakhar, 2011) by using mental (e.g., demanding cognitive activities such as counting backwards from 100 in steps of seven) or physical countermeasures (e.g., biting their tongue or moving their toes). Various experiments were conducted to assess the effects of these countermeasures on the outcomes of the CIT (e.g., Ben-Shakhar & Dolev, 1996; Elaad & Ben-Shakhar, 1991, 2009; Honts, Devitt, Winbush, & Kircher, 1996; Lykken, 1960; Peth, Suchotzki, & Gamer, 2016). These studies revealed that ANS measures are affected by both mental and physical countermeasures and could thereby enhance the false negative rates, if suspects are aware of how to use countermeasures to their advantage. Yet, no increase in false positives are expected to occur.

Leakage of Crime-related Information

Influenced by the use of mass media channels, news spreads easier than ever. Disclosure of information cannot always be prevented, yet this can alter the validity of memory detection test results. Besides news reports, crime-related information can unintentionally be leaked to suspects during their interrogation. For example, Byron Halsey came to know the location where the bodies were discovered after some guess-work. If the location of the victims would then have been used as a critical detail in the CIT, it might have triggered a false positive outcome, since the CIT effect is driven by recognition instead of actual guilt. Several studies have examined the effects of information leakage on the outcome of the CIT and although their results are not entirely consistent, it seems that leakage of information to innocent suspects may significantly increase the rate of false–positives (for a review, see Bradley, Barefoot, & Arsenault, 2011; Osugi, 2018). Thus, leakage remains a major obstacle and it can be avoided only by adopting careful police investigation practices.

In addition to avoiding leakage, using multiple relevant items, and increasing the number of questions in a CIT, more research is needed to evaluate the validity of a more specific CIT. For example, instead of the known cause of death by ‘strangulation’, it might be possible to ask a more specific question, such as, “What object was used to strangulate the victim (a) rope, b) shoe lace, c) hands, d) baton or e) wire”. This may limit the risk of information leakage as only individuals with specific knowledge of the crime in question are likely to be exposed to this information.

Agenda for Future Research

Asking the Right Questions

In order to detect a memory trace that could link a suspect to the crime, it is important to ask the proper questions in the CIT. But can experienced examiners select the appropriate crime details? Rationally, central crime items are better remembered than peripheral items and are therefore also better detectable. Research revealed higher differential SCRs for central items after a delay (Gamer et al., 2010; Nahari & Ben-Shakhar, 2011). Besides centrality, actual enactment of the crime might result in more stable memory traces (Madan & Singhal, 2012) and allow for a more accurate distinction between innocent and guilty suspects in the CIT. This effect may be moderated by the saliency of the items. Since the OR is the bodily reaction upon presentation of significant stimuli, we tend to have a larger OR to for example a picture of the victim than to a random stranger. Also in the RT-CIT, larger effects were found when stimuli that draw more attention were used (Suchotzki et al., 2017). Unfortunately, in real life the control over how a complex crime scene is perceived is beyond the control of the examiners. It cannot be assumed that all details derived from the criminal investigation are actually noticed and stored in memory, ready to be exposed in a memory detection test. Future studies could therefore focus on the link between a perpetrators’ memory and the stimuli to be tested.

Information Gathering Using the CIT

Ever since the 9/11 attacks in the United States and the increased hostility of terrorist organizations, responding to and preventing security threats has become more important than ever. Detecting potential terrorists is difficult, because in many cases the critical information (e.g., location of the planned terror attack, names of the individuals involved) is not available to the investigators. In such cases, a modified version of the CIT, labelled the searching CIT (SCIT; Osugi, 2011, 2018) has been proposed. Several studies using the SCIT with groups of individuals sharing the critical items, revealed that it has potential (Breska, Ben-Shakhar, & Gronau, 2012; Breska, Zaidenberg, Gronau, & Ben-Shakhar, 2014; Meijer, Bente, Ben-Shakhar, & Schumacher, 2013; Meijer, Smulders, & Merckelbach, 2010). It provides an opportunity to detect and reveal information that is not yet known to the investigators, which could be used to prevent malicious intents, or find a hostage. However, future research is required to further validate the SCIT and reveal the scope of its prospective usage.

Conclusion

In short, the Concealed Information Test is a method based on solid scientific principles that might resolve difficulties encountered in stress-based lie detection methods. When the appropriate items from the crime-scene are selected, and intermixed with equally plausible irrelevant options, sound results can be found using memory detection. It remains important to prevent disclosure of these items during the course of the investigation, as to prevent false positives. However, as with all forensic evidence, guilt should not be solely inferred based on the CIT. Instead, memory detection can offer scientifically valid guidance on how to proceed with suspects and thereby aid in the search for the true culprit.

References

Balloun, K. D., & Holmes, D. S. (1979). Effects of repeated examinations on the ability to detect guilt with a polygraphic examination: A laboratory experiment with a real crime. *Journal of Applied Psychology*, *64*(3), 316-322.

Ben-Shakhar, G. (1977). A further study on the dichotomization theory in detection of information. *Psychophysiology, 14*(4), 408-413.

Ben-Shakhar, G. (1991). Clinical judgment and decision-making in CQT-polygraphy. *Integrative Physiological and Behavioral Science*, *26*(3), 232-240.

Ben-Shakhar, G. (2011). Countermeasures. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and applications of the Concealed Information Test* (pp. 200-214). Cambridge, UK: Cambridge University Press.

Ben-Shakhar, G. (2012). Current research and potential applications of the concealed information test: An overview. *Frontiers in Psychology*, *3*: 342.

Ben-Shakhar, G., Bar-Hillel, M., & Lieblich, I. (1986). Trial by polygraph: Scientific and juridical issues in lie detection. *Behavioral Sciences & the Law*, *4*(4), 459-479.

Ben-Shakhar, G., & Dolev, K. (1996). Psychophysiological detection through the guilty knowledge technique: Effect of mental countermeasures. *Journal of Applied Psychology*, *81*(3), 273-281.

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.

Ben-Shakhar, G., & Furedy, J. J. (1990). *Theories and applications in the detection of deception: A psychophysiological and international perspective*. New York, NY: Springer-Verlag.

Ben-Shakhar, G., & Nahari, T. (2018) The external validity of studies examining the detection of concealed knowledge using the concealed information test. In J. P. Rosenfeld (Ed.), *Detecting Concealed Information and Deception* (pp. 59-76). Cambridge, MA: Academic Press.

Blackstone, W. (1830). *Commentaries on the Laws of England* (Vol. 2). New York, NY: Collins & Hannay.

Blakemore, S. J., Winston, J., & Frith, U. (2004). Social cognitive neuroscience: Where are we heading? *Trends in Cognitive Sciences*, *8*(5), 216-222.

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1-11.

Bradley, M. T., Barefoot, C. A., & Arsenault, A. M. (2011). Leakage of information to innocent suspects. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection:*

Theory and application of the Concealed Information Test (pp. 187-199). Cambridge, UK: Cambridge University Press.

Bradley, M. T., & Janisse, M. P. (1981). Accuracy demonstrations, threat, and the detection of deception: Cardiovascular, electrodermal, and pupillary measures. *Psychophysiology*, *18*(3), 307-315.

Bradley, M. T., & Warfield, J. F. (1984). Innocence, information, and the guilty knowledge test in the detection of deception. *Psychophysiology, 21*(6), 683-689.

Breska, A., Ben-Shakhar, G., & Gronau, N. (2012). Algorithms for detecting concealed knowledge among groups when the critical information is unavailable. *Journal of Experimental Psychology: Applied*, *18*(3), 292-300.

Breska, A., Zaidenberg, D., Gronau, N., & Ben-Shakhar, G. (2014). Psychophysiological detection of concealed information shared by groups: An empirical study of the searching CIT. *Journal of Experimental Psychology: Applied*, *20*(2), 136-146.

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*(4), 261-269.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. Yuille (Ed.), *Credibility Assessment* (pp. 51-70). Dordrecht, NL: Kluwer Academic Publishers.

Elaad, E. (1990). Detection of guilty knowledge in real–life criminal investigations. *Journal of Applied Psychology*, *75*(5), 521-529.

Elaad, E., & Ben-Shakhar, G. (1991). Effects of mental countermeasures on psychophysiological detection in the guilty knowledge test. *International Journal of Psychophysiology, 11*(2), 99-108.

Elaad, E., & Ben-Shakhar, G. (2009). Countering coutermeasures in the concealed information test using covert respiration measures. *Applied Psychophysiology and Biofeedback, 34*(3), 197-208.

Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real–life criminal guilty knowledge tests. *Journal of Applied Psychology*, *77*(5), 757-767.

Gamer, M. (2011). Detecting concealed information using autonomic measures. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 27-45). Cambridge, UK: Cambridge University Press.

Gamer, M., Kosiol, D., & Vossel, G. (2010). Strength of memory encoding affects physiological responses in the Guilty Actions Test. *Biological Psychology*, *83*(2), 101-107.

Gamer, M., Verschuere, B., Crombez, G., & Vossel, G. (2008). Combining physiological measures in the detection of concealed information. *Physiology & Behavior*, *95*(3), 333-340.

Geven, L.M., Ben-Shakhar, G., Kindt, M., & Verschuere, B. (2018). *Memory–based deception detection: Extending the cognitive signature of lying from instructed to self–initiated cheating*. Manuscript submitted for publication.

Geven, L.M., klein Selle, N., Ben-Shakhar, G., Kindt, M., & Verschuere, B. (2018). *Self-Initiated versus Instructed Cheating in the Physiological Concealed Information Test*. Unpublished manuscript, Department of Clinical Psychology, University of Amsterdam, Amsterdam, the Netherlands.

Gustafson, L. A., & Orne, M. T. (1963). Effects of heightened motivation on the detection of deception. *Journal of Applied Psychology, 47*(6), 408-411.

Gustafson, L. A., & Orne, M. T. (1965). Effects of perceived role and role success on the detection of deception. *Journal of Applied Psychology, 49*(6), 412-417.

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*(1), 84-92.

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*(3), 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*(3), 323-330.

Iacono, W. G. (1991). Can we determine the accuracy of polygraph tests? In J. R. Jennings, P. K. Ackles, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 4, pp. 201-207). London, UK: Kingsley.

Kassin, S. M., Bogart, D., & Kerner, J. (2012). Confessions that corrupt: Evidence from the DNA exoneration case files. *Psychological Science*, *23*(1), 41-45.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, E. H., & Ben-Shakhar, G. (2016). Orienting versus inhibition in the Concealed Information Test: Different cognitive processes drive different physiological measures. *Psychophysiology, 53*(4), 579-590.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, E. H., & Ben-Shakhar, G. (2017a). Unraveling the roles of orienting and inhibition in the Concealed Information Test. *Psychophysiology, 54*(4), 628-639.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, E., Nahari, T., & Ben-Shakhar, G. (2017b). Memory detection: The effects of emotional stimuli. *Biological Psychology*, *129*, 25–35.

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*(8), 605-613.

Kugelmass, S., & Lieblich, I. (1966). Effects of realistic stress and procedural interference in experimental lie detection. *Journal of Applied Psychology*, *50*(3), 211-216.

Lieblich, I., Kugelmass, S., & Ben-Shakhar, G. (1970). Efficiency of GSR detection of information as a function of stimulus set size. *Psychophysiology, 6*(5)*,* 601-608.

Lieblich, I., Naftali, G., Shmueli, J., & Kugelmass, S. (1974). Efficiency of GSR detection of information with repeated presentation of series of stimuli in two motivational states. *Journal of Applied Psychology, 59*(1), 113-115.

Luria A. (1932). *The nature of human conflicts*. New York, NY: Liveright.

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology, 43*(6), 385-388.

Lykken, D. T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology, 44*(4), 258-262.

Lykken, D. T. (1974). Psychology and the lie detector industry. *American Psychologist, 29*(10)*,* 725-739.

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*, *26*(3), 214-222.

Lykken, D. T. (1998). *A tremor in the blood: Uses and abuses of the lie detector (2nd* ed.). New York, NY: Plenum.

Madan, C. R., & Singhal, A. (2012). Using actions to enhance memory: Effects of enactment, gestures, and exercise on human memory. *Frontiers in Psychology*, *3*: 507.

Marston W. M. (1920). Reaction–time symptoms of deception. *Journal of Experimental Psychology, 3*(1), 72-87.

Meijer, E. H., Bente, G., Ben-Shakhar, G., & Schumacher, A. (2013). Detecting concealed information from groups using a dynamic questioning approach: Simultaneous skin conductance measurement and immediate feedback. *Frontiers in Psychology*, *4*: 68.

Meijer, E. H., klein Selle, N., Elber, L., & Ben-Shakhar, G. (2014). Memory detection with the Concealed Information Test: A meta-analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology*, *51*(9), 879-904.

Meijer, E. H., Smulders, F. T., & Merckelbach, H. L. (2010). Extracting concealed information from groups. *Journal of Forensic Sciences*, *55*(6), 1607-1609.

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*(2), 679-688.

Nahari, G., & Ben-Shakhar, G. (2011). Psychophysiological and behavioral measures for detecting concealed information: The role of memory for crime details. *Psychophysiology, 48(6), 733-744*.

Nahari, T., Breska, A., Elber, L., Klein Selle, N., & Ben-Shakhar, G. (2017). The external validity of the Concealed Information Test: The effect of choosing to commit a mock crime. *Applied Cognitive Psychology*, *31*(1), 81-90.

Osugi, A. (2011). Daily application of the CIT in Japan. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 253-275). Cambridge, UK: Cambridge University Press.

Osugi, A. (2018). Field findings from the Concealed Information Test in Japan. In J. P. Rosenfeld (Ed.), *Detecting Concealed Information and Deception* (pp. 97-121). Cambridge, MA: Academic Press.

Patrick, C. J., & Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, *76*(2), 229-238.

Pavlov, I. (1927). *Conditioned reflexes: An investigation of the psychological activity of the cerebral cortex*. New York, NY: Oxford University Press.

Pennebaker, J. W., & Chew, C. H. (1985). Behavioral–inhibition and electrodermal activity during deception. *Journal of Personality and Social Psychology, 49*(5), 1427-1433.

Peth, J., Suchotzki, K., & Gamer, M. (2016). Influence of countermeasures on the validity of the Concealed Information Test. *Psychophysiology, 53*(9), 1429-1440.

Peth, J., Vossel, G., & Gamer, M. (2012). Emotional arousal modulates the encoding of crime-related details and corresponding physiological responses in the Concealed Information Test. *Psychophysiology*, *49*(3), 381-390.

Podlesny, J. A., & Raskin, D. C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. *Psychophysiology*, *15*(4), 344-359.

Reid, J. E. (1945). Simulated blood pressure responses in lie-detector tests and a method for their detection. *Journal of Criminal Law and Criminology, 36*, 201-214.

Rosenfeld, J. P., Oszan, I., & Ward, A. C. (2017). P300 amplitude at Pz and N200/N300 latency at F3 differ between participants simulating suspect versus witness roles in a mock crime. *Psychophysiology, 54*(4), 640-648.

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*(1), 30-37.

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, *12*(2), 48-53.

Sokolov, E. N. (1963). *Perception and the conditional reflex*. New York, NY: Macmillan.

Suchotzki, K., Verschuere, B., Peth, J., Crombez, G., & Gamer, M. (2015). Manipulating item proportion and deception reveals crucial dissociation between behavioral, autonomic and neural indices of concealed information. *Human Brain Mapping, 36*(2), 427-439.

Suchotzki, K., Verschuere, B., Van Bockstaele, B., Ben-Shakhar, G., & Crombez, G. (2017). Lying takes time: A meta–analysis on reaction time measures of deception. *Psychological Bulletin, 143*(4), 428-453.

Trovillo, P. V. (1939). A history in lie detection. *Journal of Criminal Law and Criminology, 29*(6), 848-881.

Verschuere, B., Ben-Shakhar, G., & Meijer, E. (Eds.). (2011). *Memory detection: Theory and application of the Concealed Information Test*. Cambridge, UK: Cambridge University Press.

Verschuere, B., Crombez, G., Koster, E. H., Van Bockstaele, B., & De Clercq, A. (2007). Startling secrets: Startle eye blink modulation by concealed crime information. *Biological Psychology, 76*(1-2), 52-60.

Verschuere, B., Meijer, E., & De Clercq, A. (2011). Concealed information under stress: A test of the orienting theory in real-life police interrogations. *Legal and Criminological Psychology*, *16*(2), 348-356.

Vrij, A. (2004). Guidelines to catch a liar. In P. Granhag & L. Stromwall (Eds.), *The detection of deception in forensic contexts* (pp. 287-314). Cambridge, UK: University Press.

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

Zaitsu, W. (2016). External validity of Concealed Information Test experiment: Comparison of respiration, skin conductance, and heart rate between experimental and field card tests. *Psychophysiology*, *53*(7), 1100-1107.

View publication stats