UvA-DARE (Digital Academic Repository)

Orienting versus inhibition in the Concealed Information Test: Different cognitive processes drive different physiological measures

klein Selle, N.; Verschuere, B.; Kindt, M.; Meijers, E.; Ben-Shakhar, G.

Link to publication

Citation for published version (APA):
klein Selle, N., Verschuere, B., Kindt, M., Meijers, E., & Ben-Shakhar, G. (2016). Orienting versus inhibition in the Concealed Information Test: Different cognitive processes drive different physiological measures. *Psychophysiology*, *53*(4), 579-590. https://doi.org/10.1111/psyp.12583

Orienting versus inhibition in the Concealed Information Test: Different cognitive processes drive different physiological measures

NATHALIE KLEIN SELLE,[a,b] BRUNO VERSCHUERE,[b,c,d] MEREL KINDT,[b] EWOUT MEIJER,[c] AND GERSHON BEN-SHAKHAR[a]

[a]Department of Psychology, Hebrew University of Jerusalem, Jerusalem, Israel
[b]Department of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands
[c]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
[d]Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

Abstract

The Concealed Information Test (CIT) provides a valid tool for psychophysiological detection of concealed knowledge. However, its precise theoretical underpinnings remain a matter of debate. The differential physiological responses elicited by concealed, relevant items, relative to control items, were traditionally explained as reflecting an orienting response (OR). According to an alternative account, these responses reflect attempts to inhibit arousal. The present study examined whether and to what extent CIT detection efficiency is affected by instructions aimed at manipulating arousal inhibition (AI). One hundred and forty-eight undergraduate students completed a CIT, while electrodermal, cardiac, and respiratory measures were recorded. Half of the participants were requested to imagine that they are suspected of committing a crime and were motivated to avoid detection (presumably eliciting both OR and AI), while the other half were requested to imagine that they are witnesses of a crime and were motivated to be detected (presumably eliciting OR only). All participants were further requested to remain silent throughout the test. In both conditions, concealed items led to a similar increase in skin conductance as compared to the control items. However, the typically observed heart rate deceleration and respiratory suppression were found in suspects, but not in witnesses. These data imply that different mechanisms drive the responding of different psychophysiological measures used in the CIT, with skin conductance reflecting OR, and heart rate and respiration primarily reflecting AI.

Descriptors: Concealed Information Test (CIT), orienting response (OR), arousal inhibition (AI), skin conductance response (SCR), respiration line length (RLL), heart rate (HR)

The Concealed Information Test (CIT), first introduced by Lykken (1959) as the Guilty Knowledge Test, utilizes physiological responses to detect the presence of crime-related information in memory. In a typical CIT, examinees are presented with a series of multiple choice-like questions, each designed to examine suspects' knowledge of a distinctive crime detail assumed to be known only to individuals involved in the crime (e.g., perpetrator, victim) and the investigative authorities. For each question, this critical feature of the crime (i.e., the relevant item) is intermixed among several plausible control items (e.g., Was the murder weapon a knife? a pair of scissors? a screwdriver? an ice pick? a fork?). An innocent (unknowledgeable) suspect will not be able to differentiate between the different alternatives and is therefore expected to show similar physiological responses to all items. A guilty suspect, however, will recognize the relevant alternatives and will show differential physiological responses to these items (i.e., the CIT effect).

Orienting response (OR) theory has been the predominant approach to explain the CIT effect (e.g., Ben-Shakhar, 1977; Lieblich, Kugelmass, & Ben-Shakhar, 1970; Lykken, 1974). The OR is manifested by physiological and behavioral responses in reaction to external novel stimuli. Consequently, when a stimulus is presented repeatedly, its novelty value declines and the OR gradually habituates (Lynn, 1966; Sokolov, 1963). Importantly, an enhanced OR occurs when the stimulus holds a special significance to a subject (Sokolov, 1963). Lykken (1974) was the first to realize that this quality of ORs enables one to identify "knowledgeable suspects." Returning to our previous example, the true murder weapon (e.g., an ice pick) will be significant solely to the perpetrator and will therefore elicit a stronger OR than the other alternatives (e.g., a knife, a pair of scissors, a screwdriver, a fork). OR theory can explain many observations (see a review by Verschuere & Ben-Shakhar, 2011). First and foremost, compared to the control items, the relevant items typically elicit a pattern of enhanced responding

that characterizes OR to significant stimuli: a larger skin conductance response (SCR), a shorter respiration line length (RLL), and slower heart rate (HR; e.g., Gamer, 2011). Second, physiological responding to relevant CIT items has two important OR characteristics, namely, habituation and generalization (e.g., Balloun & Holmes, 1979; Ben-Shakhar, Frost, Gati, & Kresh, 1996).

A number of empirical findings from CIT experiments are, however, not adequately explained by OR theory. First, different CIT measures (SCR, RLL, and HR) neither correlate across subjects (Gamer, Verschuere, Crombez, & Vossel, 2008) nor within subjects across CIT questions (Gamer, Godert, Keth, Rill, & Vossel, 2008). Second, several studies have shown that HR and RLL, unlike SCR, are relatively resistant to habituation (e.g., Barry, 1977; Ben-Shakhar & Elaad, 2002; Elaad & Ben-Shakhar, 1997; Gamer, Godert et al., 2008). Finally, HR deceleration evoked by the relevant CIT items may last for 15 s, whereas, according to OR theory, HR should decelerate 1–5 s after onset of the OR-eliciting stimulus and then return to baseline (Richards & Casey, 1992). To explain this lack of convergence (i.e., fractionation) between the different physiological measures, it has been hypothesized that the different measures are driven by different mechanisms (e.g., Ambach, Stark, & Vaitl, 2011; Barry, 1996, 2006, 2009; Bradley, 2009). In other words, each physiological measure might reflect a different subprocess of the CIT.

Recognition of OR theory's weaknesses has led researchers to examine the contribution of other cognitive processes to the physiological responding in the CIT. For example, Verschuere, Crombez, Koster, Van Bockstaele, and De Clercq (2007) investigated the CIT's underlying mechanisms by using a startle eyeblink paradigm. The authors reasoned that orienting to stimuli enhances startle modulation when long lead intervals are used. Based on OR theory, they therefore anticipated enhanced startle modulation to crime pictures in comparison to control pictures. The startle data, however, showed reduced rather than enhanced startle modulation to crime pictures, leading the authors to conclude that, besides orienting, inhibition of the physiological responses also contributed to the CIT effect. Two additional experiments by Verschuere et al. (2007) supported the proposed role of inhibition. Importantly, the inhibition account can explain both the relative resistance of the cardiopulmonary measures (HR and RLL) to habituation and the prolonged HR deceleration observed in CIT studies.

Miyake et al. (2000) defined inhibition as the executive function that allows one to deliberately, and intentionally, inhibit a dominant automatic or prepotent response. Accordingly, in the CIT context, inhibition can refer to either (a) response inhibition (RI), which targets the behavioral component of the response; or (b) arousal inhibition (AI), which targets the physiological component of the response. Pennebaker and Chew (1985) further suggested that inhibition requires effort, which in turn results in physiological costs. It follows that inhibition might counterintuitively contribute to the typical physiological CIT pattern by enhancing the responses to the relevant CIT items. An influential account of deceptive behavior proposed that deception involves the inhibition of the prepotent truth response (Spence et al., 2001). Specifically, it has been assumed that the truth is activated automatically, and, consequently, individuals motivated to deceive face a conflict, which is resolved by inhibition of the truthful response. Although the CIT is not a test of deception, it was reasoned that the denial of relevant-item knowledge causes a conflict between the truthful and the required deceptive response, which is resolved by inhibiting the former response (Seymour & Schumacher, 2009; Verschuere & De Houwer, 2011). Response inhibition has been indicated to play a

role in the CIT based on reaction time (Suchotzki, Verschuere, Peth, Crombez, & Gamer, 2014) and fMRI measures (e.g., Gamer, 2014; Gamer, Bauermann, Stoeter, & Vossel, 2007; Langleben et al., 2002; Phan et al., 2005), but it remains unclear whether it plays a role in the autonomic-based CIT (see Ambach, Stark, Peper, & Vaitl, 2008a; Ambach et al., 2011). Indeed, meta-analytic studies have demonstrated that similar CIT effects were observed with the SCR measure when subjects responded deceptively to the relevant items and when they remained silent and did not give any overt responses (Ben-Shakhar & Elaad, 2003; Meijer, klein Selle, Elber, & Ben-Shakhar, 2014). Still, it can be argued that even in the silent condition, subjects motivated to conceal the relevant information implicitly inhibit the truthful response. While the role of RI is debatable, attempts to inhibit the physiological arousal associated with the relevant items in order to avoid detection characterize individuals motivated to conceal the relevant information. Consequently, the present study focuses on the AI component, and, when addressing inhibition, we refer primarily to the inhibition of experienced physiological arousal.

Indirect support for the contribution of arousal inhibition to the CIT effect comes from the emotional response modulation literature. These studies demonstrated that participants who were requested to inhibit (i.e., suppress) their emotional responses showed increased SCRs and HR deceleration, while respiration measures (respiration depth or period) were generally unaffected (e.g., Demaree et al., 2006; Gross & Levenson, 1993, 1997). Most of the response modulation research, however, focused on either experiential or expressive suppression, and only one study (Dan-Glauser & Gross, 2011) directly compared the effects of expressive and physiological suppression. The results revealed a strong overlap between the two inhibition strategies in terms of their effects on autonomic responses; in both cases, inhibition led to a decrease in cardiovascular (e.g., HR) and respiratory activity. Importantly, these findings overlap with the typically observed decrease in cardiac and respiratory activity in response to the relevant CIT items. Other research has also shown that voluntary control of cardiopulmonary measures can be successfully achieved (e.g., Bell & Schwartz, 1975; Ley, 1999; Victor, Mainardi, & Shapiro, 1978).

Some CIT studies have attempted to directly manipulate the inhibition factor (e.g., Verschuere et al., 2007). Gustafson and Orne (1965) made a first attempt by comparing the commonly used deceive condition, in which participants are motivated to avoid detection, with an additional detected condition in which participants were motivated to be detected. Participants also received feedback between two successive CIT trials and were detected significantly less with the electrodermal measure on the second trial when feedback was incompatible with their perceived roles (e.g., detected by the machine when motivated to avoid detection). Although it is likely that participants' failure on the first trial increased their motivation on the second trial, it is unclear whether this result was moderated by inhibition. Subsequent studies using motivational instructions also found little support for the role of inhibition. Horvath (1979) examined the differential effects of two opposing motivational states (i.e., motivated to be detected and motivated to avoid detection) using a card-test paradigm. Although SCR detection efficiency was higher for subjects trying to be detected, it was highly similar to that of a nonmotivated group in Horvath (1978), which suggests that SCR detection efficiency is not contingent on the need for inhibition. Zvi, Nachson, and Elaad (2012) examined the influence of the state of mind (coping or cooperative) of guilty and informed innocent participants on the outcomes of the CIT. The results indicated that both coping behavior

and guilt contributed to the enhanced SCRs to relevant information (no effect for RLL and finger pulse wave length). The enhanced SCRs were explained as reflecting a higher defensive motivation and attempts to inhibit arousal in the coping condition. A subsequent study (Elaad, 2013), which included only informed innocents, found similar results: coping instructions and an incentive for success were related to enhanced SCRs to the relevant information. Importantly, however, even in the cooperative conditions of both studies, inhibition might have played a role since all participants shared the common goal of proving their innocence. Hence, it is difficult to say which mechanism caused the differential findings.

Matsuda, Nittono, and Ogawa (2013) used a novel CIT design in which participants received two CITs: one for an item they had to conceal, and one for an item they saw being revealed to the experimenter. The authors reasoned that the disclosure of the critical item removed the need to inhibit responses to this item. Only respiration and the late positive potential showed the expected effect of concealment, while the SCR, HR, and two ERP components (N2 and P3) showed no effect. However, given the nature of the CIT, it is unclear whether the disclosure manipulation successfully eliminated all attempts to inhibit responses.

Finally, Ambach, Stark, Peper, and Vaitl (2008b) examined the effect of overt deception on the CIT's outcomes by requiring their participants to answer either deceptively or truthfully 4 s after item presentation (see also Verschuere, Crombez, Smolders, & De Clercq, 2009). Specifically, participants were motivated through the promise of a reward to perform the CIT task as well as possible. While one group was instructed to deny their knowledge and press the "no" key to all CIT items, the other group was instructed to answer truthfully and press the "yes" key to all CIT items. The authors found the CIT effect to disappear in the truth condition for the HR and RLL measures. While the SCR measure differentiated between the relevant and the control items in both conditions, the SCR effect was stronger in the lie condition than in the truth condition. However, although the delayed answering method cleverly distinguished between deceptive and orienting processes, we believe there are two key limitations to the experimental design. First, as Ambach et al. (2008b) indicated, the differential response (yes vs. no) in the two conditions might have acted as a confound and influenced the physiological responses. Second, although all participants were motivated to perform well, performing well in the truth condition was defined as answering as "correctly as possible to all questions." As the instructions only referred to the required behavioral response, there is a possibility that some participants in the truth condition might have interpreted the interrogation task as a polygraph experiment, and consequently might have attempted to inhibit their responses (see Rosenfeld, Hu, & Pederson, 2012; Verschuere, Rosenfeld, Winograd, Labkovsky, & Wiersema, 2009, for P300 studies). Thus, although some studies (see also Verschuere, Crombez, & Koster, 2004) manipulated the factor of response inhibition, they did not fully isolate this factor as different confounding variables (e.g., motivation, verbal response) were present in their experimental designs.

The foremost aim of this study was to set up a manipulation designed to isolate the OR factor by neutralizing the AI factor. This was accomplished by contrasting two conditions: a suspect condition, where participants were motivated to avoid detection, and a witness condition, where participants were motivated to be detected. Unlike previous studies, only one of the experimental conditions (i.e., suspect) was expected to create the need for inhibition. Specifically, half of the participants were requested to imagine that they are suspected of committing a crime and were motivated to successfully conceal the relevant items and avoid detection (OR and AI), while the other half was requested to imagine that they are witnesses of a crime and were motivated to successfully reveal the relevant items and to be detected (OR only). Importantly, our manipulation was directed at participants' motivation, whereby "suspects" and "witnesses" are simply labels conveying the induced motivational state. Under these conditions, witnesses welcomed the enhanced arousal elicited by the relevant items. Consequently, they oriented to the significant, recognized items but there is no reason to believe that they tried to inhibit their physiological responses. Furthermore, to avoid a possible confounding effect of overt deception, no behavioral or verbal response to the relevant items was required. It is important to emphasize that the witness condition in the present proposal differs drastically from the informed innocent condition applied in many previous CIT studies (see Bradley, Barefoot, & Arsenault, 2011, for a review). Although both informed innocents and witnesses possess the relevant information, informed innocents are typically motivated to conceal their knowledge and to appear innocent, whereas in the current study, witnesses were actually motivated to be detected.

Because AI is expected to be present only in suspects, the inhibition theory would predict larger differential responses to the relevant items in the suspect than in the witness condition. On the other hand, the OR theory would predict no difference between these two conditions, given that OR is expected in both suspects and witnesses. Based on the idea of response fractionation (Barry, 1996, 2006, 2009; Bradley, 2009), we hypothesized that the various physiological measures will be differentially affected by our experimental manipulation. Specifically, we expected to find an interaction between the experimental conditions (suspect vs. witness) and the physiological measures, such that the SCR, which is considered as a pure OR index (e.g., Barry, 1996, 2006, 2009) will be unaffected by the manipulation, while both RLL and HR will lose some of their discriminative power in the witness condition where no inhibition is called for.

Method

Participants

One hundred and forty-eight undergraduate students (94 women) of the Hebrew University of Jerusalem with an age range of 18–36 ($M = 24.2$ $SD = 2.9$ years) participated in this study. Most of the participants were psychology students, and all of them were native speakers of Hebrew and received either course credits or an average payment of 35 NIS (equivalent to approximately 10 USD) for their participation. Each participant read and signed a consent form indicating that participation was voluntary and that they could withdraw from the experiment at any time without penalty.

Data Acquisition and Reduction

The experiment was conducted in an air-conditioned laboratory. The apparatus included a HP Compaq DC 5800 Microtower computer that was used to control stimulus presentation and compute skin conductance, respiration, and heart rate.

Electrodermal activity was recorded using a constant voltage system (0.5 V ASR Atlas Researches, Hod Hasharon, Israel), two Ag/AgCl electrodes (0.8 cm diameter) filled with a 0.05 M NaCL electrolyte (TD-246, Discount Disposables) and an A/D (NB-MIO-12) converter with a sampling rate of 50 Hz. Electrodes were placed on the distal phalanges of the left index and left ring finger. SCRs were defined as the maximal increase in conductance obtained from 1–5 s after stimulus onset.

The electrocardiogram (ECG) was recorded by placing three Ag/AgCl electrodes filled with an electrode paste in a standard Einthoven lead I configuration: one electrode attached to the distal phalange of the left index finger (i.e., one of the SCR electrodes), one electrode attached to the right wrist, and the ground electrode attached to the left wrist. The ECG signal was sampled at 500 Hz, digitized at 12-bit resolution, and filtered using a band-pass of 0.5–35 Hz. MATLAB was used to detect the R peaks, calculate the distance between them, and apply a semiautomatic artifact detection and rejection procedure (similar to, e.g., De Clercq, Verschuere, De Vlieger, & Crombez, 2006). Prior to analysis, the interbeat intervals were converted to HR in beats per minute (bpm) per real-time epoch (1 s). These second-by-second poststimulus HR values were baseline-corrected by subtracting the average HR value in the 3 s preceding stimulus onset (i.e., the prestimulus baseline value), resulting in 15 poststimulus difference scores (ΔHR). The average of all ΔHR scores has been found to outperform the minimum of all ΔHR scores as a detection measure (Gamer, Verschuere et al., 2008) and was therefore the preferred statistic when analyzing the data.

Respiration was recorded using a pneumatic tube positioned around the thoracic area. Respiration responses were defined on the basis of the total RLL, which is a composite measure of respiratory amplitude (depth of breathing) and respiratory cycle (rate of breathing), during the 13-s interval following stimulus onset. Following Elaad, Ginton, and Jungman (1992), we defined each response as the mean of 10 length measures (0.1 s after stimulus onset through 13.1 s after stimulus onset, 0.2 s through 13.2 s after stimulus onset, etc.). In other words, 10 13-s windows were created, each beginning 0.1 s later than the previous window, and the RLL was defined as the mean of the 10 length measures computed for the 10 windows.

For all three measures, individual responses were removed if excessive movements were made during the measurement window or if the response was an outlier (Z score larger than 5 or smaller than −5). Further, participants whose standard deviation of raw SCR scores was below 0.01 μS[1] were considered to be skin conductance nonresponders (approximately 25% of the normal population does not show measurable SCRs to orienting stimuli, as reported by Venables & Mitchell, 1996), and their SCR data was eliminated from all analyses.

Due to individual differences in physiological responsiveness and baseline activation, within-subject standard scores were calculated for each physiological channel separately (Ben-Shakhar, 1985). Further, considering that several physiological measures (especially SCR) are sensitive to habituation, the standard scores were computed for each item within a block of three questions. Thus, the within-subject standard scores were computed by subtracting the mean of all responses to both control and relevant items within each block of questions (buffer and catch items were excluded from the standardization) from each individual raw response and dividing this difference by the respective standard deviation (see Ben-Shakhar & Elaad, 2002; Elaad & Ben-Shakhar, 1997). Each block contained a total of 30 items (6 relevant and 24 control items), and consequently the Z scores in this study were computed relative to the mean and standard deviation of the examinees' responses to these 30 items. A detection score was created by averaging the Z scores of all 12 relevant items for each participant.

---

1. Rather than determining nonresponsivity based on the average SCR (e.g., Venables & Mitchell, 1996), we chose the variance criterion for two reasons: (1) it eliminates individuals who show no variability in their responses across items; and (2) it does not eliminate individuals who display small responses to most items, but may still show differential responses to the relevant items.

Procedure

All participants came to the laboratory and were allocated to either the suspect or the witness condition. Experimenter 1 read the same fictional newspaper article (see Appendix A), describing a burglary, to both suspects and witnesses. The article familiarized participants with all relevant items that were used in the subsequent CIT. Five versions of the article were used such that each version described a different set of six relevant items (see Appendix B for a description of the relevant sets). Except for the variation in relevant items, all five versions were identical and were randomly assigned to participants, such that each of the five item-sets served as the relevant set for 20% of the participants in each condition. The four other sets of items served as control items in the CIT. In this manner, each item served both as a relevant and as a control item across subjects, thereby eliminating potential biasing effects.

Participants were asked to envision the described situation and identify with their assigned role: suspects were told to imagine that they executed the burglary and witnesses were told to imagine that they saw the burglary happen. After the examiner had read out the article, all participants were requested to take a few minutes to carefully read the article by themselves. In order to prevent differences in explicit memory between the two groups, all participants received a short questionnaire testing their knowledge of the critical items before continuing to the CIT. Experimenter 2, who was unaware of the specific item set used, attached the SCR and HR electrodes as well as the RLL band and conducted the CIT examination.

The CIT. Participants in both conditions were reminded about their assigned roles (suspect vs. witness) and informed of the basic principles of the CIT. More specifically, participants were informed that their physiological responses would change when they recognize the information related to the crime. Participants imitating suspects were motivated to avoid detection and promised a bonus of 10 NIS (about 3 USD) as an incentive for a successful concealment of the relevant crime items (i.e., come out as innocent). Participants imitating witnesses were motivated to be detected and were promised a bonus of 10 NIS as an incentive to help the examiners identify the relevant crime details. The bonus was paid when the average SCR Z score, computed across all relevant items in the two blocks (i.e., SCR detection score), was below 0.1 for suspects or above 0.1 for witnesses. All participants were further requested to remain silent throughout the entire CIT.

Following an initial rest period of 2 min, participants were presented with six CIT questions targeting different features of the crime described in the newspaper article. The order of these questions was randomly determined. To maintain participants' attention, a short break was inserted after three questions. Each question was presented on the computer monitor for 10 s and at the same time played through the computer's loudspeakers. Following question presentation, the different items appeared for 5 s each, with an interstimulus interval of 10–14 s (Breska, Maoz, & Ben-Shakhar, 2011). The first item was always a neutral, buffer item designed to absorb the initial orienting response. Next, two repetitions of six items were presented in a random order: the relevant item, four control items, and one catch item. Catch items (see Appendix B) also referred to as target items in the CIT literature, were included as an extra means of assuring that participants' attention remained focused on the items presented (see also Verschuere, Crombez, Degroote, & Rosseel, 2010). Each catch item was denoted under the question itself, and participants were instructed to press the left mouse button as quickly as possible when identifying the catch item among the presented answer

items. No further button presses were required. In sum, participants were presented with 6 questions $\times$ 13 items (1 buffer, 2 relevant, 8 control, and 2 catch items), totaling 78 items.

Following the CIT, a multiple-choice recognition memory test was administered to examine whether participants recognized the relevant items. The memory test consisted of the same six questions used in the CIT; however, all possible answer items were presented simultaneously. Following this memory check, participants were asked to rate the valence, significance, and arousal level of all six relevant items and six randomly chosen control items (one from each question). These ratings were obtained to examine whether the relevant items carried the same importance for participants imitating suspects as for participants imitating witnesses. Then, participants were asked to rate their level of motivation during the experiment and, although not instructed to, whether and what kind of countermeasures they applied. Finally, all participants were debriefed and compensated for their participation in the experiment.

Data Analysis

As concealed information is associated with cardiac and respiratory suppression, the RLL and HR Z scores were multiplied by -1 prior to analysis. A preliminary Condition (suspect vs. witness) $\times$ Item Type (relevant vs. control) $\times$ Poststimulus Second (1 through 15) mixed three-way analysis of variance (ANOVA) was performed on the raw HR data to obtain a clear view of how the heart rate pattern develops with time. For the main analysis, a mixed 2×3 ANOVA, with condition (suspect vs. witness) as a between-subjects factor and response type (SCR, RLL, and HR) as a within-subject factor, was performed on the standardized physiological data. This was followed by two sets of orthogonal planned contrasts, which were designed to examine more closely the main effect of response type and its interaction with condition. The first main effect contrast compared SCR to the average RLL and HR and the second main effect contrast compared RLL and HR. The second set of (interaction) contrasts followed a similar pattern, but compared the physiological response differences between suspects and witnesses. A rejection region of $p < .05$ was used for all statistical tests, and Cohen's f values were computed as effect size estimates (Cohen, 1988). According to Cohen (1988), the values of $f = 0.1$, $f = 0.25$, and $f = 0.40$ correspond to small, medium, and large effects, respectively.

In addition, CIT detection efficiency was evaluated using measures derived from signal detection theory (SDT; e.g., Green & Swets, 1966; Swets, Tanner, & Birdsall, 1961). SDT is particularly useful for analyzing psychophysiological detection data, and it has been applied extensively in the area of memory detection (e.g., Ben-Shakhar & Elaad, 2003; National Research Council, 2003). SDT measures are typically derived by comparing the detection score distribution of knowledgeable (guilty) individuals (in this case, either suspects or witnesses) with the detection score distribution of unknowledgeable (innocent) individuals. Two measures of detection efficiency can be derived from this comparison: first, Cohen's d, which expresses the standardized difference between the means of two detection score distributions, where d of 0.20, 0.50, and 0.80 are considered as small, moderate, and large effects, respectively (Cohen, 1988); and second, the area under the receiver operating characteristic (ROC) curve, which describes the detection efficiency of the CIT across all possible cutoff points on the detection score and varies between 0 and 1, with a chance level of 0.5.

As a sample of unknowledgeable (innocent) participants was not included in the present study, the expected detection score distribution among unknowledgeable participants was simulated using the method proposed by Meijer, Smulders, Johnston, and Merckelbach (2007). This simulation procedure depends on the assumption that the relevant items have no special meaning for unknowledgeable individuals, and, therefore, there is no reason to expect that they would elicit systematically different responses. Since all responses were standardized within individuals, the responses of unknowledgeable participants to the individual relevant items have a mean of zero and a unit standard deviation. Consequently, the average of these standardized responses will have a mean of zero and a standard deviation of 1 divided by the square root of the number of questions or repetitions used for generating the detection score. In the present experiment, the detection scores were computed as an average of 12 relevant items (six relevant items repeated twice). In addition, it is assumed that these responses have a normal distribution, and, therefore, the simulated detection score distribution among innocents was created by taking random samples n (either $n_{suspects}$ or $n_{witnesses}$) from a normal distribution; $N(0, 1/\sqrt{12})$. Then, we compared these simulated distributions with the empirical detection score distributions obtained for the suspects and witnesses and computed Cohen's d, and the area under the ROC curve. This process was repeated 10,000 times for each physiological measure in each condition, and the mean d value and area, as well as the 95% confidence interval (CI) of the area across the 10,000 repetitions, were computed.

The justification of this simulation method was recently examined by Meijer et al. (2014). These researchers applied simulations on 16 data sets from studies that used samples of unknowledgeable participants and compared the outcomes of the simulations with those reported in the original studies. The comparisons showed no systematic differences between the outcomes of the two methods, thus providing empirical validation for the use of the simulations to estimate the responses of unknowledgeable examinees.

Results

Data of one participant were excluded from analysis because of noncompliance with the instructions, and skin conductance data of 30 participants (20%) were removed due to nonresponsivity. Furthermore, heart rate data of eight participants and memory data of the first 56 participants were lost due to technical issues. We started to obtain the subjective arousal, valence, and significance ratings from the 26th participant and lost the data of two participants due to technical issues; therefore, there are 27 missing values for these ratings. Thus, while sample size varies between 92–139 for the different analyses, all analyses of the three physiological measures were based on data of at least 109 participants. For the remaining participants, 4% of all SCRs, 1% of all RLL, and 4% of all HR responses to the individual stimuli were removed due to excessive movements and outliers.

Memory, Countermeasures, and Subjective Responses

A marginally significant difference in the number of correctly remembered relevant items (scores ranging from 0–6) was found between the suspect ($M = 5.87$) and witness conditions ($M = 5.98$), $t(90) = -1.99$, $p = .05$, $d = -.42$; nonetheless, participants from both conditions remembered details nearly perfectly.

Thinking "exciting" thoughts to specific items was the preferred countermeasure across all subjects (suspects: 27%, witnesses: 26%). Another frequently used countermeasure was the intentional distortion of respiratory rate or depth (suspects: 12%, witnesses: 15%). Finally, some subjects silently recited the items in their head (suspects: 6%, witnesses: 16%).

Table 1. *Mean Subjective Rating Scores of the Relevant and Control Items in Suspects and Witnesses*

Category	N	Mean (SD) relevant items	Mean (SD) control items
Significance			
Suspect	60	7.20 (1.40)	2.43 (1.44)
Witness	61	7.47 (1.74)	2.18 (1.40)
Valence			
Suspect	60	3.98 (1.49)	4.72 (1.49)
Witness	61	5.08 (1.64)	4.16 (1.44)
Arousal			
Suspect	60	6.67 (1.71)	3.06 (1.50)
Witness	61	6.68 (1.89)	2.87 (1.23)

Table 1 shows the mean subjective rating scores (significance, valence, and arousal) of the relevant and control items in each condition. Each of the subjective rating scores was submitted to a Condition (suspect vs. witness) × Item Type (relevant vs. control) mixed measures ANOVA. A main effect of item type was observed for both the significance, $F(1,119) = 586.05$, $f = 1.55$, $p < .001$, and arousal, $F(1,119) = 364.58$, $f = 1.23$, $p < .001$, ratings, indicating higher significance and arousal level for the relevant items compared to the control items. No main effects of condition (significance: $F(1,119) = .00$, $p = .955$; arousal: $F(1,119) = .18$, $p = .671$) and no interaction effects were found for these ratings (significance: $F(1,119) = 1.60$, $p = .208$; arousal: $F(1,119) = .28$, $p = .597$). For the valence ratings, a significant interaction effect of condition and item type was observed, $F(1,119) = 19.51$, $f = .28$, $p < .001$. Follow-up paired sample t tests revealed that suspects significantly rated the relevant items as less pleasant than the control items, $t(59) = -2.78$, $p = .007$, $d = -.50$, whereas witnesses rated the relevant items as more pleasant than the control items, $t(60) = 3.47$, $p = .001$, $d = .60$.

Physiological Measures

A preliminary 2 × 2 × 15 mixed three-way ANOVA was performed on the second-by-second raw ΔHR data, with condition (suspect vs. witness) serving as a between-subjects factor and both item type (relevant vs. control) and poststimulus second (1–15) as within-subject factors. Most interestingly, after correcting for sphericity ($\varepsilon = 0.38$), a significant interaction between the factors condition, item type, and poststimulus second was found, $F(14,1918) = 5.29$, $f = .12$, $p < .001$. This interaction was further examined by performing, within each experimental condition, a 2 × 15 ANOVA with item

type and poststimulus second as within-subject factors. In the suspect condition, in addition to two significant main effects for item type, $F(1,67) = 58.70$, $f = .17$, $p < .001$, and poststimulus second, $F(14,938) = 46.90$, $f = .56$, $p < .001$, a significant interaction between item type and poststimulus second was observed, $F(14,938) = 11.09$, $f = .26$, $p < .001$, indicating larger HR deceleration to the relevant than to the control items. In the witness condition, on the other hand, no statistically significant item type effect was observed, $F(1,70) = .12$, $p = .73$, but a significant main effect of poststimulus second, $F(14,980) = 26.37$, $f = .41$, $p < .001$, and a significant interaction between item type and poststimulus second were observed, $F(14,980) = 2.67$, $f = .10$, $p = .02$. Looking at Figure 1B, there seems to be a difference in HR between the relevant and control items for seconds 6–9 (i.e., larger deceleration to relevant than to control items) and for seconds 13–15 (i.e., larger deceleration to control than to relevant items). Follow-up paired sample t tests on both the 6–9 s period, $t(70) = -1.55$, $p = .125$, $d = -.24$, and the 13–15 s period, $t(70) = 1.83$, $p = .071$, $d = .35$, were not statistically significant. These results reveal that, although the relevant and control items showed a somewhat different trend in the ΔHR data, they produced similar HR deceleration across the entire 15 s (see Figure 1A,B).

A mixed 2 × 3 ANOVA with condition (witness vs. suspect) as a between-subjects factor and response type (SCR vs. RLL vs. HR) as a within-subject factor was performed to analyze the standardized responses to relevant items (see Figure 2). After correcting for sphericity ($\varepsilon = .78$), the mixed measures ANOVA revealed a significant main effect of condition, $F(1,107) = 62.15$, $f = .43$, $p < .001$, indicating larger average relative responses to the relevant items in the suspect than in the witness condition. In addition, a significant main effect of response type, $F(2,214) = 39.05$, $f = .48$, $p < .001$, along with a significant interaction of Condition × Response Type were found, $F(2,214) = 11.45$, $f = .25$, $p < .001$.

The main effect of response type and its interaction with condition were further examined using two sets of planned contrasts. The first main effect contrast revealed a statistically significant difference between SCRs and cardiopulmonary responses (RLL and HR) across the two conditions, $t(214) = 7.62$, $f = .59$, $p < .001$, indicating that the SCR measure was associated with larger relative responses to the relevant items, as compared with the other two measures. The second main effect contrast revealed no statistically significant difference between the RLL and HR responses across conditions, $t(214) = -1.22$, $p = .223$. The first interaction contrast revealed that the difference between SCR and the mean of RLL and HR responses was significantly larger in the witness than in the
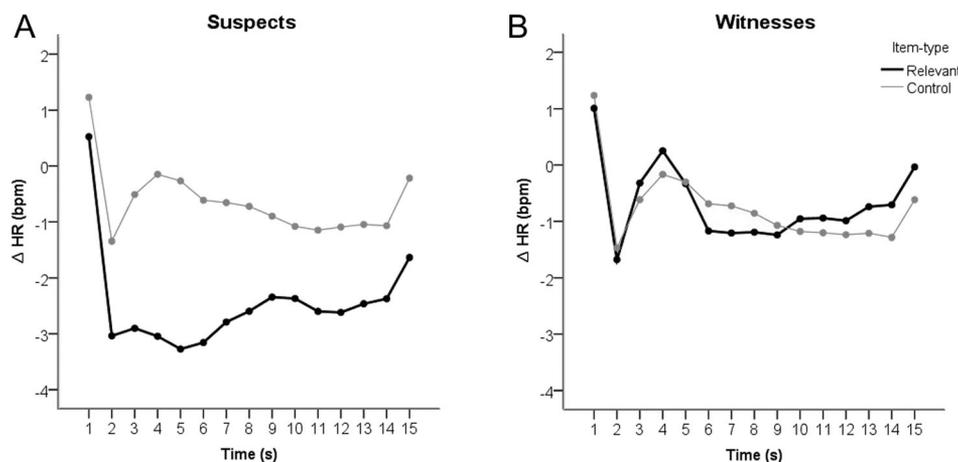
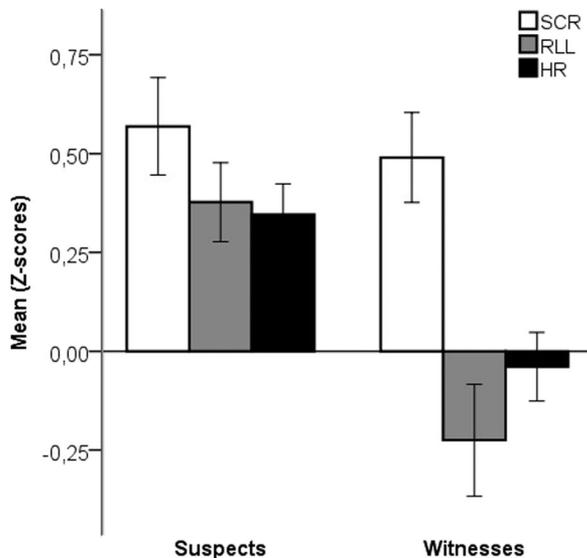**Figure 1.** HR change to relevant and control items of suspects (A) and witnesses (B).

Figure 2. Standardized mean skin conductance response, respiration line length, and heart rate to relevant items of suspects and witnesses.

suspect condition, $t(214) = -3.80$, $f = .29$, $p < .001$. This result reflects the fact that witnesses did not display the typically observed HR and RLL suppression in response to the relevant items. An additional interaction contrast revealed a significant difference in RLL versus HR responses between conditions, $t(214) = 1.72$, $f = .09$, $p = .003$, reflecting the fact that the RLL average Z score was larger than the HR among suspects, but lower among witnesses.

Signal detection analysis. In order to analyze detection efficiency of classifying individuals as knowledgeable (guilty) versus unknowledgeable (innocent), we compared the detection score distribution of each physiological measure in each condition with the expected detection score distribution among simulated innocents. For each of the six comparisons, we computed Cohen's d and the ROC area as well as their respective 95% confidence intervals. These results, which are displayed in Table 2, reveal that detection efficiency with the SCR measure was significantly larger than a chance area of 0.50 for both suspects and witnesses, with almost identical ROC areas in the two conditions (0.85 and 0.83, respectively). In contrast, both RLL and HR produced a statistically significant ROC area only for suspects. A comparison of the areas obtained in the suspect and the witness conditions revealed a statistically significant difference for the RLL, $Z = 5.98$, $p < .001$, and the HR, $Z = 4.38$, $p < .001$, but not for the SCR, $Z = .35$, $p = .37$. Unexpectedly, the ROC area for the RLL measure in the witness condition was significantly smaller than a chance area of 0.50 (i.e., larger RLLs were elicited by the relevant items, compared to the neutral items).[2] Importantly, similar to the responses elicited by relevant items in the witness condition, the SCR increased, the RLL lengthened, and the HR of witnesses did not change in response to catch items (see Table 2); for suspects, the HR slightly increased in response to catch items, $t(67) = 2.44$, $p = .017$, $d = 0.30$.

---

2. It is important to note that the RLL lengthening remained after disqualifying participants who indicated using respiratory countermeasures during the CIT ($M_{all} = -0.21$, $M_{noCountermeasures} = -0.12$). The difference between the average RLL of all witnesses and the average RLL of witnesses who used no countermeasures was not statistically significant, $t(61) = 1.49$, $p = 0.142$, $d = 0.17$.

Discussion

To advance the theoretical underpinnings of the CIT, the present study investigated whether different cognitive mechanisms drive the responding of different psychophysiological measures. By contrasting two opposing motivations, using the classic suspect condition (where participants were motivated to avoid detection) and our novel witness condition (where participants were motivated to be detected), we assessed the roles of orienting and inhibition processes in the physiological responses to relevant crime details. The typically observed CIT response pattern (Gamer, 2011)—namely, an increased SCR, a shorter RLL, and HR deceleration—was replicated as expected in the suspect condition, but not in the witness condition. Specifically, our motivational manipulation did not influence CIT detection efficiency with the SCR measure, but it did affect detection efficiency when using cardiopulmonary (i.e., HR and RLL) measures, showing that the SCR measure reflects orienting processes, while both the RLL and HR primarily reflect inhibition processes. This was reflected by a statistically significant interaction between experimental condition and physiological response type and by the ROC results, supporting the multiple mechanisms account. The main findings were supported by our catch item results: in both conditions, these items elicited similar responses (as reflected by the three physiological measures) to those elicited by the relevant items in the witness condition. Since catch items had to be identified by all participants, they are expected to elicit an OR, but not inhibition.

We did not find a difference between conditions with the SCR measure. This may raise the question of whether our statistical test had sufficient power. To clarify this question, we conducted a power analysis, which indicated that, to obtain a statistical power of at least 0.80 for detecting a medium effect size (i.e., Cohen's d of 0.50), the sample size in each condition should be about 55. The sample size in both our experimental conditions exceeded this number ($n_{suspects} = 57$, $n_{witnesses} = 60$). In addition to the power analysis, we computed the JZS Bayes factor, which is a numerical value quantifying how well the null and alternative hypothesis predicts the data (Rouder, Speckman, Sun, Morey, & Iverson, 2009). The JZS Bayes factor was evaluated using Jeffrey's (1961) criteria: values smaller than 1, between 1 and 3, and between 3 and 10, respectively, designate "no evidence," "anecdotal evidence," and "substantial evidence" for the null hypothesis. We found a JZS Bayes factor of 4.51, meaning that there is substantial evidence for the null hypothesis.

Unexpectedly, we observed a lengthening of the RLL in the witness condition. This lengthening may have resulted from attempts made by participants in the witness condition to alter their physiological responses. Such deliberate actions are typically used by guilty suspects to avoid detection (i.e., countermeasures), but when motivated to be detected, participants may try to facilitate detection by creating enhanced responding to the relevant items. However, this account is questionable for the following two reasons: First, the RLL has been found to be more resistant to both mental and physical countermeasures than the SCR (Ben-Shakhar & Dolev, 1996; Honts, Devitt, Winbush, & Kircher, 1996). Second, the RLL lengthening remained after disqualifying participants who indicated having used respiratory countermeasures during the CIT. Because this is the first study to observe such RLL lengthening, we think that a replication will be crucial to examine whether this unexpected finding is robust.

With regard to the HR measure, we did not observe the often-reported biphasic HR pattern (i.e., an initial increase followed by a decrease). In contrast to most CIT studies that require an overt verbal no response, no overt response was required in the present experiment.

Table 2. *Standardized Means (SDs) of the Physiological Responses to Relevant and Catch Items in Suspects and Witnesses; Effect Size (d) and Area Under the ROC Curve (a) with 95% CIs*

Category	N	Mean Z score (SD) relevant items	Mean Z score (SD) catch items	Effect size (Cohen's d) with 95% CI	Area under the ROC curve (a) with 95% CI
SCR					
Suspect	57	0.54 (0.44)	1.07 (2.20)	1.47 (1.05, 1.89)	0.85 (0.81, 0.89)
Witness	60	0.50 (0.42)	0.97 (1.38)	1.39 (0.98, 1.79)	0.83 (0.79; 0.87)
RLL					
Suspect	73	0.35 (0.39)	−0.23 (0.44)	1.03 (0.68, 1.38)	0.76 (0.71, 0.80)
Witness	74	−0.21 (0.55)	−0.19 (0.42)	−0.48 (−0.80, −0.15)	0.39 (0.34, 0.44)
HR					
Suspect	68	0.32 (0.29)	−0.10 (0.33)	1.11 (0.75, 1.48)	0.78 (0.72, 0.84)
Witness	71	−0.01 (0.33)	0.01 (0.32)	−0.04 (−0.37, 0.30)	0.51 (0.44, 0.57)

Note. The *d* and *a* values were computed by comparing the detection score distribution of either suspects or witnesses with a simulated distribution of innocents.

Our results therefore support the hypothesis raised by Verschuere, Crombez et al. (2009) that verbalization accounts for the initial acceleration. In line with our findings, Verschuere, Crombez et al. (2009) observed an immediate deceleration of the HR when participants remained silent. Furthermore, in accordance with previous CIT studies, the HR deceleration lasted for 15 s following stimulus onset. OR theory predicts, as noted before, that the HR only decelerates in the initial 5 s and then returns to baseline (Richards & Casey, 1992), thus providing an additional support for the inhibition account of HR.

Importantly, as heightened significance is related to larger ORs, it is important to stress that there was no difference between the two conditions in the perceived significance level of the relevant items. This means that for both groups the critical information carried the same weight or, in other words, the importance of relevant items was just as high for suspects as for witnesses. Hence, the observed RLL and HR differences cannot be explained by differences in orienting.

Response Fractionation

Clearly, the current results are inconsistent with Sokolov's (1963) classical concept of a unitary OR and seem to support the notion of response fractionation (Barry, 1996, 2006, 2009; Bradley, 2009). According to the classical view, the different physiological measures covary in response to manipulations of novelty, intensity, and significance. However, the accumulation of evidence from OR studies gradually revealed that correlations between the different OR components are low and that these components are affected by different factors. More specifically, only the SCR showed the effects of novelty, intensity, and significance expected from Sokolov's theory. Bernstein, Taylor, Weinstein, and Reidel (1985) correspondingly showed that manipulating the significance of presented tones increased the SCR but did not affect the HR. The present findings show a similar fractionation between the SCR and cardiopulmonary measures (RLL and HR). In particular, only the SCR showed the expected significance effect in the witness condition where the inhibition factor was eliminated, and thus the OR factor was isolated. This strongly suggests that the RLL and HR measures are driven by inhibition.

Inhibition and the CIT

Comparisons between the current observations and previous findings on inhibition utilizing the CIT reveal both similarities and differences. Ambach et al. (2008b) used a highly similar design to the one used in the present study, and, when comparing the results across studies, there are important similarities in findings. Interestingly, no CIT effect was found with the respiratory and heart rate measures in

the present witness condition, which is consistent with the results observed in the truthful condition of Ambach et al. (2008b). Taken together, both studies seem to support the notion that the HR and RLL measures are underlaid by inhibition. Further, while Verschuere et al. (2007) found smaller HR deceleration in the condition that aimed to remove inhibition, Matsuda et al. (2013) did not.

Regarding the SCR measure, unlike the results reported by Ambach et al. (2008b), we did observe a strong CIT effect in the witness condition with the SCR measure, which was virtually identical to the effect observed in the suspect condition. As discussed by Ambach et al. (2008b), the differential SCR response in the truth condition might have also been the result of the greater physiological impact and the lower frequency of the yes answers compared to the no answers. In the present study, it was therefore crucial to equalize verbal answer types across conditions. Our SCR findings are also compatible with data of a recent study by Suchotzki et al. (2014). Participants in this study were instructed to actively deny knowledge of one crime and admit knowledge of a second crime. Although overt denial was crucial for finding a CIT effect with both RT and fMRI measures, it was not necessary for the SCR.

Applied Implications

A recent meta-analysis of CIT studies (Meijer et al., 2014) found a similar average effect size in studies that required a deceptive verbal response (i.e., no) and in studies where no verbal response was required. However, these studies relied only on the SCR measure, and, unfortunately, sufficient data for other physiological measures were unavailable. As SCR reflects orientation to the significant, recognized items, it is not surprising that it is unaffected by a deceptive verbal response. However, since our results indicate that both RLL and HR are driven by AI, they might also be sensitive to manipulations of RI and consequently to a deceptive response. In particular, larger CIT detection efficiency is expected with these measures when examinees give deceptive answers to the relevant items. While this issue definitely warrants further research, we believe it is safe to recommend that practitioners instruct their examinees to give deceptive (i.e., no) answers when the items are presented to them.

Limitations of Current Study

First, although the obtained significance ratings indicate that the importance of the relevant items was identical for witnesses and suspects, it cannot be excluded that factors other than inhibition contributed to the differences that we observed between the two conditions. The valence ratings, for instance, indicated that the crime-relevant

items carried different meanings for participants imitating suspects and witnesses. Further, there may be expected differences in deception, feelings of guilt, and possibly shame between the two conditions. Future studies could deal with these limitations by, for example, rating deceptive participants' experienced level of felt emotions. This will allow for an examination of whether increased ratings on these factors are related to the enhanced responses to relevant items. However, since the basic CIT effect for the RLL and HR completely vanished in witnesses, which has not been found in other CIT studies investigating deception (e.g., deceptive verbal response—Meijer et al., 2014; informed innocence—Bradley, MacLaren, & Carle, 1996; Gamer, Godert et al., 2008), it is unlikely that this factor can fully explain the present findings.

Second, as noted previously, as this is the first CIT study to observe lengthening of the RLL, we believe that a replication will be crucial to examine whether this surprising finding is robust. The third limitation refers to the possibility that witnesses' increased RLL influenced their HR, which is known to synchronize with respiratory rhythm (e.g., Yasuma & Hayano, 2004). In particular, the increased RLL of witnesses might have partly masked a decelerative HR response.

Finally, all participants were checked for memory of the relevant items before continuing to the CIT, which clearly differs from real-life CIT examinations. However, for the present experimental manipulation, the need to prevent differences in explicit memory between the conditions outweighed the need to create a more realistic laboratory environment.

Conclusions

In summary, using a novel manipulation contrasting a suspect condition (where participants were motivated to avoid detection) with a witness condition (where participants were motivated to be detected), our research showed that different mechanisms underlie the different autonomic CIT measures. Although the SCR was not affected by our manipulation, the typical shortening of the RLL and slowing of the HR were solely observed in participants imitating suspects. While orientation to the significant items occurred in both conditions, attempts at arousal inhibition characterized only the suspect condition. We accordingly conclude that the SCR measure reflects orienting processes, while both the RLL and HR primarily reflect inhibition processes.

The current study provides valuable information on how the different physiological measures detect crime-relevant information. As we pointed out, the differential responses to the relevant items have so far been accounted for by single underlying mechanisms—primarily by either orienting or inhibition. The data presented here, however, warrant a different approach. We therefore hope that the present study will encourage future research and that the accumulation of findings from these studies will strengthen and broaden current CIT theory. A strong theory is not only the pinnacle of any scientific method, but might also encourage the use of the CIT as an applied investigative tool.

Appendix A

Note. The face of the accomplice has been made invisible as real mugshots were used in the experiment.

Appendix B

	City in which the burglary took place	Time of the burglary	The homeowner	The point of entry	Type of jewellery stolen	Face of the accomplice
Set 1	Tel Aviv	0:00	Minister of defense	Back gate	Diamond ring	Picture 1
Set 2	Jerusalem	1:00	Head of the mossad	Bedroom balcony	Gemstone bracelet	Picture 2
Set 3	Haifa	2:00	Chief of air force	Kitchen window	Pearl necklace	Picture 3
Set 4	Beer Sheva	3:00	Chief of army	Front door	Emerald earrings	Picture 4
Set 5	Eilat	4:00	Chief of police	Roof	Golden watch	Picture 5
Catch items	Ashdod	5:00	Head of the shabak	Maintenance room	Silver brooch	Picture 6

<p style="text-align:center">CIT question categories</p>

References

Ambach, W., Stark, R., Peper, M., & Vaitl, D. (2008a). An interfering go/no-go task does not affect accuracy in a Concealed Information Test. *International Journal of Psychophysiology*, *68*, 6–16. doi: 10.1016/j.ijpsycho.2007.11.004

Ambach, W., Stark, R., Peper, M., & Vaitl, D. (2008b). Separating deceptive and orienting components in a Concealed Information Test. *International Journal of Psychophysiology*, *70*, 95–104. doi: 10.1016/j.ijpsycho.2008.07.002

Ambach, W., Stark, R., & Vaitl, D. (2011). An interfering *n*-back task facilitates the detection of concealed information with EDA but impedes it with cardiopulmonary physiology. *International Journal of Psychophysiology*, *80*, 217–226. doi: 10.1016/j.ijpsycho.2011.03.010

Balloun, K. D., & Holmes, D. C. (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*, 316–322. doi: 10.1037/0021 9010.64.3.316

Barry, R. J. (1977). The effect of "significance" upon indices of Sokolov's orienting response: A new conceptualisation to replace the OR. *Physiological Psychology*, *5*, 209–214. doi: 10.3758/BF03335318

Barry, R. J. (1996). Preliminary process theory: Towards an integrated account of the psychophysiology of cognitive processes. *Acta Neurobioliae Experimentalis*, *56*, 469–484.

Barry, R. J. (2006). Promise versus reality in relation to the unitary orienting reflex: A case study examining the role of theory in psychophysiology. *International Journal of Psychophysiology*, *62*, 353–366. doi: 10.1016/j.ijpsycho.2006.01.004

Barry, R. J. (2009). Habituation of the orienting reflex and the development of preliminary process theory. *Neurobiology of Learning and Memory*, *92*, 235–242. doi: 10.1016/j.nlm.2008.07.007

Bell, I. R., & Schwartz, G. E. (1975). Voluntary control and reactivity of human heart rate. *Psychophysiology*, *12*, 339–349. doi: 10.1111/j.1469-8986.1975.tb01302.x

Ben-Shakhar, G. (1977). A further study on the dichotomization theory in detection of information. *Psychophysiology*, *14*, 408–41. doi: 10.1111/j.14698986.1977.tb02974.x

Ben-Shakhar, G. (1985). Standardization within individuals: A simple method to neutralize individual differences in skin conductance. *Psychophysiology*, *22*, 292–299. doi: 10.1111/j.1469-8986.1985.tb01603.x

Ben-Shakhar, G., & Dolev, K. (1996). Psychophysiological detection through the guilty knowledge technique: Effects of mental countermeasures. *Journal of Applied Psychology*, *81*, 273–281. doi: 10.1037/0021-9010.81.3.273

Ben-Shakhar, G., & Elaad, E. (2002). Effects of questions' repetition and variation on the efficiency of the guilty knowledge test: A reexamination. *Journal of Applied Psychology*, *87*, 972–977. doi: 10.1037/0021-9010.87.5.972

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*, 131–151. doi: 10.1037/00219010.88.1.131

Ben-Shakhar, G., Frost, R., Gati, I., & Kresh, Y. (1996). Is an apple a fruit? Semantic relatedness as reflected by psychophysiological responsivity. *Psychophysiology*, *33*, 671–679. doi: 10.1111/j.1469-8986.1996.tb02363.x

Bernstein, A. S., Taylor, K. W., Weinstein, E., & Reidel, J. (1985). The effect of stimulus significance on relatively sustained (tonic-like) and relatively transient (phasic-like) aspects of electrodermal, heart rate, and eyeblink response. *Biological Psychology*, *21*, 183–228. doi: 10.1016/0301-0511(85)90029-8

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*, 1–11. doi: 10.1111/j.1469-8986.2008.00702.x

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. doi: 10.1017/CBO9780511975196.011

Bradley, M. T., MacLaren, V. V., & Carle, S. B. (1996). Deception and nondeception in guilty knowledge and guilty action polygraph tests. *Journal of Applied Psychology*, *81*, 153–160. doi: 10.1037/0021-9010.81.2.153

Breska, A., Maoz, K., & Ben-Shakhar, G. (2011). Inter-stimulus intervals for skin conductance response measurement. *Psychophysiology*, *48*, 437–440. doi: 10.1111/j.1469-8986.2010.01084.x

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

Dan-Glauser, E. S., & Gross, J. J. (2011). The temporal dynamics of two response-focused forms of emotion regulation: Experiential, expressive, and autonomic consequences. *Psychophysiology*, *48*, 1309–1322. doi: 10.1111/j.1469-8986.2011.01191.x

De Clercq, A., Verschuere, B., De Vlieger, P., & Crombez, G. (2006). Psychophysiological analysis (PSPHA): A modular script-based program for analyzing psychophysiological data. *Behavior Research Methods*, *38*, 504–510. doi: 10.3758/BF03192805

Demaree, H. A., Schmeichel, B. J., Robinson, J. L., Pu, J., Everhart, D. E., & Berntson, G. G. (2006). Up- and downregulating facial disgust: Affective, vagal, sympathetic, and respiratory consequences. *Biological Psychology*, *71*, 90–99. doi: 10.1016/j.biopsycho.2005.02.006

Elaad, E. (2013). Effects of goal-and task-oriented motivation in the guilty action test. *International Journal of Psychophysiology*, *88*, 82–90. doi: 10.1016/j.ijpsycho.2013.02.004

Elaad, E., & Ben-Shakhar, G. (1997). Effects of item repetitions and variations on the efficiency of the guilty knowledge test. *Psychophysiology*, *34*, 587–596. doi: 10.1111/j.1469-8986.1997.tb01745.x

Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real-life criminal guilty knowledge tests. *Journal of Applied Psychology*, *77*, 757–767. doi: 10.1037/0021-9010.77.5.757

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. doi: 10.1017/CBO9780511975196.003

Gamer, M. (2014). Mind reading using neuroimaging: Is this the future of deception detection? *European Psychologist*, 19, 172–183. doi: 10.1027/1016-9040/a000193

Gamer, M., Bauermann, T., Stoeter, P., & Vossel, G. (2007). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping*, 28, 1287–1301. doi: 10.1002/hbm.20343

Gamer, M., Godert, H. W., Keth, A., Rill, H-G., & Vossel, G. (2008). Electrodermal and phasic heart rate responses in the guilty actions test: Comparing guilty examinees to informed and uninformed innocents. *International Journal of Psychophysiology*, 69, 61–68. doi: 10.1016/j.ijpsycho.2008.03.001

Gamer, M., Verschuere, B., Crombez, G., & Vossel, G. (2008). Combining physiological measures in the detection of concealed information. *Physiology and Behavior*, 95, 333–340. doi: 10.1016/j.physbeh.2008.06.011

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: John Wiley & Sons.

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, 64, 970–986. doi: 10.1037//0022-3514.64.6.970

Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, 106, 95–103. doi: 10.1037/0021-843X.106.1.95

Gustafson, L. A., & Orne, M. T. (1965).The effects of perceived role and role success on the detection of deception. *Journal of Applied Psychology*, 49, 412–417. doi: 10.1037/h0022798

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. doi: 10.1111/j.1469-8986.1996.tb02111.x

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. doi: 10.1037/0021-9010.63.3.338

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. doi: 10.1037/0021-9010.64.3.323

Jeffreys, H. (1961). *Theory of probability*. London, UK: Oxford University Press.

Langleben, D. D., Schroeder, L., Maldjian, J., Gur, R. C., McDonald, S., Ragland, J. D., . . . Childress, A. R. (2002). Brain activity during simulated deception: An event-related functional magnetic resonance study. *NeuroImage*, 15, 727–732. doi: 10.1006/nimg.2001.1003

Ley, R. (1999). The modification of breathing behavior: Pavlovian and operant control in emotion and cognition. *Behavior Modification*, 23, 441–479. doi: 10.1177/0145445599233006

Lieblich, I., Kugelmass, S., & Ben Shakhar, G. (1970). Efficiency of GSR detection of information as a function of stimulus set size. *Psychophysiology*, 6, 601–608. doi: 10.1111/j.1469-8986.1970.tb02249.x

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, 43, 385–388. doi: 10.1037/h0046060

Lykken, D. T. (1974). Psychology and the lie detector industry. *American Psychologist*, 29, 725–739. doi: 10.1037/h0037441

Lynn, R. (1966). *Attention, arousal and the orientation reaction*. Oxford, UK: Pergamon Press.

Matsuda, I., Nittono, H., & Ogawa, T. (2013). Identifying concealment-related responses in the concealed information test. *Psychophysiology*, 50, 617–626. doi: 10.1111/psyp.12046

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, 879–904. doi: 10.1111/psyp.12239

Meijer, E. H., Smulders, F. T. Y., Johnston, J. E., & Merckelbach, H. L. G. J. (2007). Combining skin conductance and forced choice in the detection of concealed information. *Psychophysiology*, 44, 814–822. doi: 10.1111/j.1469-8986.2007.00543.x

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex '"frontal lobe"' tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100. doi: 10.1006/cogp.1999.0734

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 Academies Press.

Pennebaker, J. W., & Chew, C. H. (1985). Behavioral-inhibition and electrodermal activity during deception. *Journal of Personality and Social Psychology* 49, 1427–1433. doi: 10.1037/0022-3514.49.5.1427

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. doi: 10.1016/j.acra.2004.11.023

Richards, J. E., & Casey, B. J. (1992). Development of sustained visual attention in the human infant. In B. A. Campbell, H. Hayne, and R. Richardson (Eds.), *Attention and information processing in infants and adults: Perspectives from human and animal research* (pp. 30–60). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rosenfeld, J. P., Hu, X., & Pederson, K. (2012). Deception awareness improves P300 based deception detection in concealed information tests. *International Journal of Psychophysiology*, 86, 114–121. doi: 10.1016/j.ijpsycho.2012.06.007

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t*-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16, 225–237. doi: 10.3758/PBR.16.2.225

Seymour, T. L., & Schumacher, E. H. (2009). Electromyographic evidence for response conflict in the exclude recognition task. *Cognitive, Affective, & Behavioral Neuroscience*, 9, 71–82. doi: 10.3758/CABN.9.1.71

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

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*, 12, 2849–2853. doi: 10.1097/00001756-200109170-00019

Suchotzki, K., Verschuere, B., Peth, J., Crombez, G., & Gamer, M. (2014). Manipulating item proportion and deception reveals crucial dissociation between behavioral, autonomic and neural indices of concealed information. *Human Brain Mapping*, 36, 427–39. doi: 10.1002/hbm.22637

Swets, J. A., Tanner, W. P., Jr., & Birdsall, T. C. (1961). Decision processes in perception. *Psychological Review*, 68, 301–340. doi: 10.1037/h0040547

Venables, P. H., & Mitchell, D. A. (1996). The effects of age, sex and time of testing on skin conductance activity. *Biological Psychology*, 43, 87–101. doi: 10.1016/0301-0511(96)05183-6

Verschuere, B., & Ben-Shakhar, G. (2011). Theory of the concealed information test. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 128–148). Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511975196.008

Verschuere, B., Crombez, G., Degrootte, T., & Rosseel, Y. (2010). Detecting concealed information with reaction times: Validity and comparison with the polygraph. *Applied Cognitive Psychology*, 24, 991–1002. doi: 10.1002/acp.1601

Verschuere, B., Crombez, G., & Koster, E. (2004). Orienting to guilty knowledge. *Cognition & Emotion*, 18, 265–279. doi: 10.1080/02699930341000095

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, 52–60. doi: 10.1016/j.biopsycho.2007.06.001

Verschuere, B., Crombez, G., Smolders, L., & De Clercq, A. (2009). Differentiating orienting and defensive responses to concealed information: The role of verbalization. *Applied Psychophysiology and Biofeedback*, 34, 237–244. doi: 10.1007/s10484-009-9093-2

Verschuere, B., & De Houwer, J. (2011). Detecting concealed information in less than a second: Response latency-based measures. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 46–63). Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511975196.004

Verschuere, B., Rosenfeld, J. P, Winograd, M. R., Labkovsky, E, & Wiersema, R. (2009). The role of deception in P300 memory detection. *Legal and Criminological Psychology*, 14, 253–262. doi: 10.1348/135532508X384184

Victor, R., Mainardi, J., & Shapiro, D. (1978). Effects of biofeedback and voluntary control procedures on heart rate and perception of pain during the cold pressor test. *Psychosomatic Medicine*, *40*, 216–226. doi: 10.1097/00006842-197805000-00004

Yasuma, F., & Hayano, J. (2004) Respiratory sinus arrhythmia: Why does the heartbeat synchronize with respiratory rhythm? *Chest*, *125*, 683–690. doi: 10.1378/chest.125.2.683

Zvi, L., Nachson, I., & Elaad, E. (2012). Effects of coping and cooperative instructions on guilty and informed innocents' physiological responses to concealed information. *International Journal of Psychophysiology*, *84*, 140–148. doi: 10.1016/j.ijpsycho.2012.01.022