Distinguishing true from false confessions using physiological patterns of concealed information recognition – A proof of concept study

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

DOI
10.1016/j.biopsycho.2020.107902

Publication date
2020

Document Version
Final published version

Published in
Biological Psychology

Biological Psychology 154 (2020) 107902

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Distinguishing true from false confessions using physiological patterns of concealed information recognition – A proof of concept study

Linda Marjoleine Geven[a,b,*], Gershon Ben-Shakhar[b], Saul Kassin[c], Bruno Verschuere[a]

[a] Department of Clinical Psychology, University of Amsterdam, The Netherlands
[b] Department of Psychology, Hebrew University of Jerusalem, Israel
[c] Department of Psychology, John Jay College of Criminal Justice, United States

ARTICLE INFO

Keywords:
Memory detection
Psychophysiology
Confessions
Interrogations
Police interviews

ABSTRACT

Wrongful conviction cases indicate that not all confessors are guilty. However, there is currently no validated method to assess the veracity of confessions. In this preregistered study, we evaluate whether a new application of the Concealed Information Test (CIT) is a potentially valid method to make a distinction between true and false admissions of guilt. Eighty-three participants completed problem-solving tasks, individually and in pairs. Unbeknownst to the participants, their team-member was a confederate, tempting the participant to break the experimental rules by assisting during an individual assignment. Irrespective of actual rule-breaking behavior, all participants were accused of cheating and interrogated. True confessors but not false confessors showed recognition of answers obtained by cheating in the individual task, as evidenced by larger physiological responses to the correct than to plausible but incorrect answers. These findings encourage further investigation on the use of memory detection to discriminate true from false confessions.

In Bruton v. United States, 1968, the United States supreme court declared that a defendants' confession "*is probably the most probative and damaging evidence that can be admitted against him*" (p. 8). Indeed, people find it difficult to believe that anyone would confess to a crime they did not commit (Kassin & Wrightsman, 1980, 1981). A confession can thereby have an overwhelming impact on jurors and judges. Research has indicated that even when no further evidence linking a suspect to a crime was presented, the mere presence of a confession tripled the chance to be found guilty at trial rather than to be acquitted (Leo & Ofshe, 1998). This may explain why false confessions contribute to almost 29 % of the cases investigated by the Innocence Project (www.innocenceproject.org; see also Drizin & Leo, 2004; Garrett, 2011). The lasting detrimental effects of wrongful convictions on all involved parties highlight the need for independent measures to verify the veracity of a confession.

Given the profound influence of confession evidence on trial outcomes, all confessions should be reviewed with caution to determine their veracity. To do so, it is important to look at personal characteristics of the suspect (e.g., dispositional risk factors such as youth or cognitive impairment; Drizin & Leo, 2004; Gudjonsson, 2003) and situational factors (e.g., confrontational interrogation tactics such as the presentation of false evidence and minimization; Kassin, 2015) that might increase the probability that an innocent suspect signs a confession. Yet, even when mock-jurors read transcripts of a murder trial in which the defendants' confession was highly coerced and ruled as inadmissible by the judge, they still did not fully discount the confession when reaching a verdict (Kassin & Sukel, 1997). This pattern is found even when the confession was reported secondhand by an informant who was motivated to implicate the defendant (Neuschatz, Lawson, Swanner, Meissner, & Neuschatz, 2008; Wetmore, Neuschatz, & Gronlund, 2014). Knowledge of possible risk factors might therefore not be sufficient for laypeople to discriminate true from false confessions.

However, even for trained investigators, district attorneys, and judges it is challenging to distinguish true from false confessions. In a study by Kassin, Meissner, and Norwick (2005), prison inmates provided self-incriminating statements that were either true or false. The recorded statements were observed by police investigators who judged the veracity of the confessions. Although they were highly confident in their decision, the overall accuracy rate (53.9 %) was not significantly different from a random guess. Likewise, in a replication study specifically targeting confessions made by juvenile delinquents, the percentage of correct judgements remained at chance level (52.8 %; Honts, Kassin, & Craig, 2014; see also Honts, Forrest, & Stepanescu, 2019). Indeed, false confessions seem indistinguishable from true confessions both in terms of the details provided (Garrett, 2010) and other content

* Corresponding author at: Department of Criminology, Leiden University.
E-mail address: l.m.geven@law.leidenuniv.nl (L.M. Geven).

https://doi.org/10.1016/j.biopsycho.2020.107902
Received 15 October 2019; Received in revised form 29 April 2020; Accepted 29 April 2020
Available online 18 May 2020

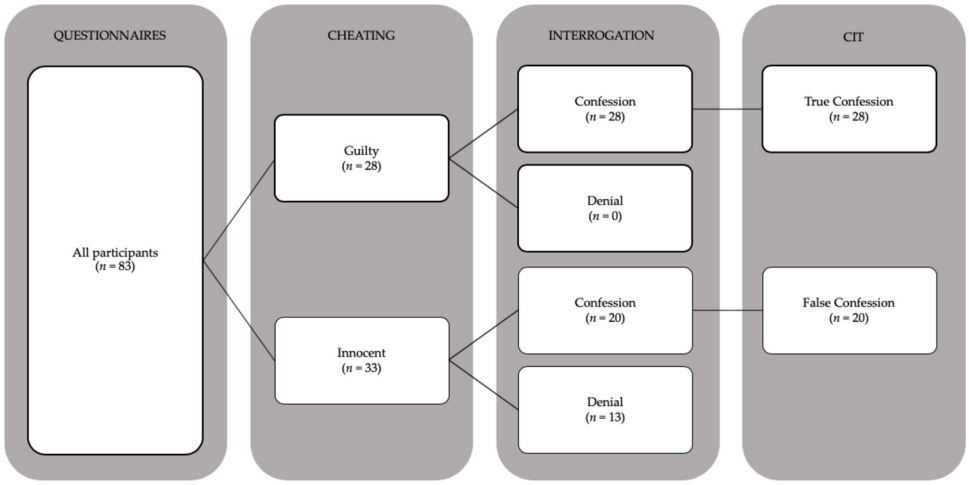

Fig. 1. Flow of the experimental procedure creating true and false confessions.
Note. The comparison of interest concerns the CIT-scores of the true and false confessors.

cues (e.g., expressions of remorse; Appleby, Hasel, & Kassin, 2013).

While false confessions represent a minority in all criminal cases handled by the courts, the consequences can be detrimental when they are not correctly recognized as such. False admissions of guilt have been shown to taint eyewitness identifications (Hasel & Kassin, 2009), statements from alibi witnesses (Marion, Kukucka, Collins, Kassin, & Burke, 2016), forensic experts (Kassin, Dror, & Kukucka, 2013), and contribute to tunnel vision amongst investigators. Moreover, if the main focus of the criminal investigation locks in on an innocent suspect, the true perpetrator escapes capture, remaining free to commit other crimes (e.g., the Central Park jogger case; Burns, McMahon, & Burns, 2012). Meanwhile, wrongful imprisonment can stigmatize the innocent (Clow & Leach, 2015) and trigger severe mental health issues (Grounds, 2005; Scott, 2010).

In the current study, the applicability of the Concealed Information Test (CIT; Lykken, 1959; Verschuere et al., 2011) to evaluate the veracity of confessions is explored. A key distinction between guilty and innocent suspects lies in the involvement in the crime and physical presence at the scene. The CIT, assessing the recognition of these intimate details derived from the investigation of the criminal act, targeting perpetrators' knowledge, may thereby distinguish true from false admissions of guilt.

The CIT examination involves the presentation of several questions, each followed by one true detail of the crime and several plausible control items, for example 'Where was the victim's body found? a) bathroom, b) kitchen, c) bedroom, d) garden, e) living room'. During the sequential presentation of all items, psychophysiological responses –most commonly skin conductance, heart rate, and respiration– are measured. Innocent examinees, who are unaware of factual information about the location of the victim in the house, are expected to respond similarly to all presented alternatives. However, recognizing salient crime details in the knowledgeable examinee will be apparent from an increase of the skin conductance (i.e., orienting response; Lykken, 1974; Sokolov, 1963), a deceleration of the heart rate and suppression of the respiration cycle (i.e., arousal inhibition; klein Selle, Verschuere, Kindt, Meijer, & Ben-Shakhar, 2016; klein Selle, Verschuere, Kindt, Meijer, & Ben-Shakhar, 2017).

The CIT was introduced by David Lykken, who reasoned that physiological responses should be used not to detect lying in itself, but rather to verify the presence or absence of crime-related details in the memory of the suspect. In the first study (Lykken, 1959), 49 participants were randomly assigned to commit mock-crimes or remain innocent. In the subsequent examination, all participants were asked questions targeting intimate details of the crimes, that would only be familiar to those actually involved. The comparison of skin conductance responsivity upon presentation of critical items to those elicited by control items revealed a high classification rate (nearly 95 %), providing initial evidence for the validity of the CIT. A more recent meta-analysis (Meijer, klein Selle, Elber, & Ben-Shakhar, 2014) supports the validity of multiple physiological measures to correctly detect both presence or absence of intimate crime knowledge, showing large effects for the skin conductance, heart rate and respiration measure (Cohen's *d* of 1.55, 0.89, and 1.11, respectively).

In the current study, we investigated whether true confessions can be distinguished from false confessions using psychophysiological memory detection. We used a variation of the procedure described by Russano, Meissner, Narchet, and Kassin (2005) to elicit true and false confessions in a laboratory environment. Similar deception paradigms were used by Exline, Thibaut, Hickey, and Gumpert (1970). Participants were paired with a research confederate and faced with various problem-solving tasks of which some had to be solved individually, while others as a pair. During the trivia quiz that had to be solved in an individual manner, the research confederate actively sought help from the participant. Participants who broke the experimental rule by assisting the female confederate with her set of questions and giving her the correct answers, were thus guilty of cheating; those who merely solved their own set of questions were deemed innocent. Independent of actual guilt, all participants were accused of cheating and interrogated by the experimenter in an accusatory manner. Both true and false confessors were examined using the CIT, to verify whether they showed recognition of the answers that they could have solved for the confederate (see Fig. 1). We expected that true confessors would exhibit differential physiological responses to the critical items in comparison to control items, while false confessors would show similar responding to all items, indicating non-recognition of the confederates' answers.

1. Method

Ethical approval was obtained from the Ethic Review Board of the University of Amsterdam and archived under number 2018-CP-9071. All participants provided written informed consent before taking part in the study, stating that study participation was voluntary and the experiment could be terminated at any time without consequences. This study was preregistered on https://osf.io/pbjt5. Task scripts, data and other materials are publicly available on https://osf.io/9dk5g.

1.1. Participants

Participants were recruited through a university portal and received course credits or a monetary compensation. Participants were required

to be fluent in Dutch and between the age of 18 and 40. The initial sample (i.e., before exclusion; see below) consisted of 83 individuals (79.5 % female), and participants were 21.86 years old (SD_{age} = 3.73, range from 18 to 38) on average.

In the current paradigm, a confederate requested participants to help with two trivia questions while the pair was instructed to solve them individually. Based upon their performance on the trivia quiz with the confederate, the participants were classified as guilty (i.e., those solving the confederates' individual trivia questions and thereby violating the experimental rule), or innocents (i.e., those who refused to help the confederate and worked independently during the individual tasks), see Fig. 1. Twenty-two participants were excluded due to several reasons: Eighteen participants only knew the answer to one of the two cheating questions; two participants initially agreed to assist the confederate, but later declined (due to knowledge contamination, these participants were also excluded from analyses); and two participants exercised their right to withdraw participation and their experimental session was immediately terminated and their data were excluded from analyses.

The remaining sample consisted of 28 individuals who were guilty of helping the confederate and thereby breaking the experimental rule (M_{age} = 21.68, SD_{age} = 4.26), and 33 innocents (i.e., those who refused to help the confederate and worked independently during the individual tasks, M_{age} = 21.73, SD_{age} = 3.45).

Regardless of their actual cheating behavior, all participants were accused of cheating and interrogated by the experimenter. All 28 participants who cheated on the individual task admitted the transgression and confessed (i.e., true confessors). Of the 33 innocents, 13 participants denied helping the confederate (39.4 %; i.e., true deniers; M_{age} = 23.08, SD_{age} = 3.66) and twenty participants gave a coerced false confession (60.6 %; i.e., false confessors; M_{age} = 20.85, SD_{age} = 3.08). Our preregistered comparison of interest concerns the contrast between the 28 true confessors versus the 20 false confessors. There were no significant differences between true and false confessors in age, $t(46)$ = 0.74, p = .462, d = 0.22, 95 % CI [-0.36;0.79], BF_{01} = 2.75, or gender, X^2 (1) = 0.87, p = .351, ϕ_c = 0.14, BF_{01} = 1.78.

1.2. Procedure

Participants were recruited for research on the influence of personality on problem-solving, disguising the true aim of the experiment. The experiment was divided into five distinct parts: personality questionnaires (for exploratory analyses), problem-solving assignments (allowing to engage in rule-breaking behavior), the interrogation (aimed at eliciting a confession), the CIT (aimed at distinguishing true from false confessions) and an elaborate debriefing.

1.2.1. Personality questionnaires

Before participating in the study, participants received the link to an online form to complete the 60-item HEXACO Personality Inventory Revised (60-HEXACO-PI-R; Ashton & Lee, 2009) and the Gudjonsson Compliance Scale (GCS; Gudjonsson, 1989), followed by demographical information on age, gender and field of study. At least seven days from completing the online questionnaires, participants were invited to the experimental session in the lab. Exploratory analyses using these measures are provided in Supplementary Appendix A.

1.2.2. Problem-solving assignments

Upon arrival, participants were met by two female research assistants who alternated the roles of experimenter and confederate posing as another participant. Participants were asked to leave all personal belongings in the adjacent room and wash their hands in preparation of the physiological measurements in a later stage of the study. Then, the pair was escorted into a small testing room with blinded windows and asked to read the information brochure and sign the informed consent.

Generally following the paradigm described by Russano et al. (2005), the experimenter provided each participant with a booklet including three individual problem-solving assignments followed by three group tasks. The experimenter instructed that it was important for the results to comply with the instructions on whether the assignment should be solved individually or in pairs, thereby implying that cooperating on the individual tasks was not allowed. After instructing participants to set a timer with an alarm after five minutes on the computer for each of the six tasks, the experimenter left the room and the participants started completing their own problem-solving booklet.

During the last individual task, the confederate feigned difficulties with answering two trivia quiz questions presented in her booklet. Approximately three minutes into the task, the confederate stated they were probably not allowed to work together, but nevertheless asked whether the participant would be willing to solve two questions. The confederate requested help without actually showing the participants the questions until the participant agreed to provide help. Upon agreement of the participant –thereby breaking the critical experimental rule– the confederate showed her booklet and the critical questions to the participant (i.e., 'Which social media logo is shown here?' and 'Which continent hosted the Winter Olympics this year?'[1]), and asked for the correct answers (i.e., Snapchat and Asia). If the participant declined helping the confederate after a second request, the confederate did not attempt to seek further information from the participant.

After participants finished all problem-solving tasks, or after the final timer, the experimenter returned to the room and collected the booklets in order to review correctness of the answers. In fact, the experimenter did not check the answers at this stage, as to remain blind on whether participants had actually cheated. After roughly two minutes, the experimenter returned to the participants and said that she wanted to talk to both participants separately, starting with the confederate. During that time, the participants waited alone for five minutes in the small experimental room.

1.2.3. Interrogation

The interrogation protocol and physical layout of the room were conducted following the guidelines recommended by Inbau, Reid, Buckley, and Jayne (2013). The protocol was designed by the first author, who completed training and is certified in the Reid technique. Although this coercive technique has been criticized (see Gudjonsson, 2003), we deliberately used the guidelines in an attempt to provoke a large number of false confessions.

While the confederate remained in the adjacent room, participants sat in the small experimental room, equipped with two chairs, a desk and bare walls. The participant was always seated in the far corner of the room, so that the experimenter sat between the participant and the door during the interrogation. The experimenter–blind to actual cheating behavior– entered the room and directly accused the participant of not following instructions by working together on an individual task. The experimenter stated that the cheating behavior posed a problem as the data were intended to be published in a scientific journal. Lastly, the experimenter said that she should probably inform her supervisor about this problem but was reluctant to do so as this would bear unknown consequences.

The experimenter then sat down next to the participant and adopted a more open attitude to offer the deal. She pretended to have come up

[1] The Olympic Winter games were held in South Korea, in February 2018. During piloting in March 2018, participants readily knew the continent in which the Winter Olympics were held. During the final stages of data collection (July 2018), many participants seemed to have difficulty answering the question and had to be excluded from data analysis (n = 18). We therefore changed the question to 'Which continent has, besides the largest surface, the most inhabitants?' in September 2018. Please note that the correct answer of both questions remains Asia and no changes had to be made to the CIT.

with a possible solution, in which the participant would make a new appointment to re-do the experiment with a new partner. This way, the data of the current session could be overwritten and the supervisor did not have to be alerted. Participants were informed that they would receive participation credits after the new appointment. While holding this monologue, the experimenter wrote down a confession statement to be signed by the participant who accepted the deal. The statement read 'I –name of participant– confess that I worked together with another participant on a task that had to be solved individually'. During the entire interrogation phase, participants' objection and denials were interrupted. Minimization techniques were also actively deployed to elicit a confession, specifically targeting prosocial behavior and absence of malicious intent (e.g., 'I would also help another person if they would ask me to' and 'I am sure you didn't know that it would influence the results'). When these tactics did not result in a signed confession, the experimenter pretended to look for the supervisor and returned to the participant stating that the supervisor was lecturing for 30 more minutes. If the participant expressed the desire to wait for the supervisor (and did not confess), the experimenter aborted the interrogation and continued to the CIT.

1.2.4. CIT

In this part of the study, the experimenter attached the respiration belts as well as the skin conductance and heart rate electrodes and conducted the full CIT procedure (due to hardware failure, the RLL signal could not be analyzed and is not reported). Participants were told that they were suspected of cheating and instructed to prove their innocence in the subsequent polygraph test. To mimic the situation in which an unknowledgeable outcome would be favorable for the suspect, we offered participants a €5 (or an equivalent in participation credits) incentive for successful concealment of their knowledge of the correct answers. The bonus was paid to all participants regardless of their actual CIT results.

Following an initial rest period of two minutes after attachment of the electrodes, participants were presented with the CIT questions and answering alternatives. The alternative answers to the two questions the participants could have cheated on during the trivia quiz were presented on the screen either as an image or as a written word (see Fig. 2 and Supplementary Appendix B for items). Beside the respective presentation on the computer screen, participants heard the spoken version through their headphones. Participants were asked to respond with a verbal 'no' to each alternative. The order of the questions and their answering alternatives was randomly determined. Each alternative answer was either presented as an image or as a written word. A self-paced break was given after the first four questions to maintain participant's attention. The audio files were pre-recorded by a third party who was blind to the procedure. The question remained on the screen for 10 s, followed by the answering alternatives that were presented for 5 s each, with a mean inter-stimulus interval of 18 s (range from 16 to 20). Between item presentation, a fixation-cross appeared on the screen to maintain the attention of the participant.

The first answering alternative following the question was always a buffer item designed to absorb the initial orienting response. Subsequently, the critical item, four control items and a single catch item were presented in a random order. Catch items were included in the CIT to ensure the participants' attention to all presented items. Upon identifying the catch item amongst the presented alternatives, participants were instructed to repeat this specific item verbally. In the current CIT, the catch items consisted of random numbers ranging from 1 to 10, either in written words or as an image. Besides from repeating the catch items, participants were instructed to respond to all other items with a verbal "no". Altogether, participants were presented with two blocks of four questions, each consisting of seven items (i.e., 1 buffer, 1 critical, 4 control, and 1 catch item), totaling 56 items. For a visual representation of the CIT, see Fig. 2.

After the CIT, the participants performed an autobiographical Implicit Association Test (aIAT; Sartori, Agosta, Zogmaister, Ferrara, & Castiello, 2008) for exploratory research purposes (results not reported here).

Recall and recognition memory tests were administered to examine whether participants correctly remembered the critical items from the trivia quiz. First, the two quiz questions for which the confederate asked the participant for help were sequentially presented on the computer screen with a text balloon in which participants could freely enter the answer. All participants were informed of the option to enter 'I don't know'. Answers were coded as either correct (1) or incorrect (0), leading to a total possible score of 2, reflecting perfect recall. Then, both cheating questions were counterbalanced to be depicted twice in a forced-choice format, with five answer options (i.e., the critical and four control items) being presented once in written words and once with images. Participants were asked to select the most appropriate option if they did not know the correct answer. Answers were coded as either correct (1) or incorrect (0), leading to a recognition scoring range from $0-4$. Then, participants rated their general experience of being accused of cheating on a 7-point Likert scale ranging from 1 (*very negative*) to 7 (*very positive*). Also, they were asked to rate the amount of stress they had experienced during the interrogation, on a 7-point Likert scale ranging from 1 (*not at all stressful*) to 7 (*very stressful)*. Then, participants rated six questions designed to assess their motivational state on a 5-point Likert scale. This questionnaire measured how well participants were able to focus on the screen during the CIT, how involved they were in the study and how much they tried to avoid detection and appear innocent on the CIT. Finally, participants rated their effort to suppress or raise their physiological responses during the test and freely elaborated on whether they had used strategies to avoid detection.

1.2.5. Debriefing

All participants were asked what they thought the study was about. From there, the experimenter fully described the procedure and the true nature of the study. The act of misleading the participants to reach the goal of the study was explicitly mentioned and explained. Participants were told that the set-up of the study was specifically designed to aid cheating behavior, adding that cheating was an admirable prosocial act in this case. Furthermore, participants were informed that no consequences would follow and that the supervisor was in fact not alerted. Also, the risk and dangers of falsely confessing to an act not committed were explicitly discussed with participants. Participants received an extensive written debriefing, explaining the relevance of the study and how minimizing interrogation techniques may lead to wrongful convictions. At last, participants were asked not to discuss the real aim or procedure of the study with others.

1.3. Data acquisition and reduction

The experiment was conducted in an air-conditioned laboratory. Item presentation was performed using Presentation® software (Version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). Psychophysiological responses were measured and recorded with Vsrrp89 software, developed by the Technical Support Social and Behavioral Sciences at the University of Amsterdam.

Electrodermal activity was recorded with an amplifier using a 50 Hz, sine-shaped excitation voltage with an amplitude of 1Vpp. Two curved-shape sintered silver-silver chloride (Ag/AgCl) electrodes (20 × 16 mm) were connected to the palmar surface of the distal phalanges of the left index and left ring finger with adhesive tape. The Skin Conductance Response (SCR) was measured from 1 s to 5 s after item onset and defined as the maximal increase in conductance during this time window.

The ECG measure was acquired by placing a set of three Ag/AgCl electrodes (3M™ Red Dot™ disposables, type 2249-50) in a standard Einthoven lead-II configuration: one electrode attached near the distal

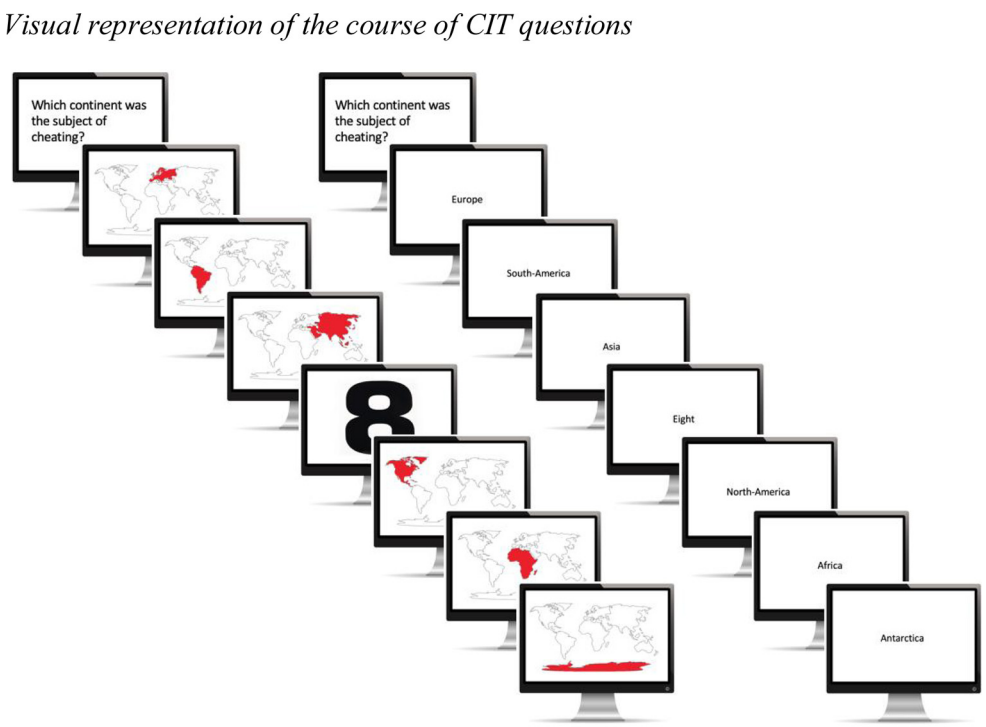

Fig. 2. Visual representation of the course of CIT questions.
Note. True confessors, but not false confessors, were hypothesized to show physiological responses indicating recognition of the critical item (e.g., Asia) in comparison to control items (e.g., South-America, North-America, Africa, and Antarctica).

end of the left collarbone, one electrode placed near the distal end of the right collarbone, and one electrode placed on the left lateral base of the chest. Prior to analysis, the inter-beat intervals were converted to Heart Rate (HR) in beats per minute per real-time epoch (1 s). The 15 second-by-second post-item HR values were baseline-corrected by subtracting the average pre-item baseline HR value (mean HR in the three seconds preceding item onset), resulting in 15 post-item difference scores (ΔHR). The average of these 15 scores was used as the HR deceleration dependent measure.

To eliminate individual differences in raw response patterns, we used within-subject standardized scores for each physiological measure. For each participant we computed the response to the critical items relative to the mean and the standard deviation of the total response distribution within each block of two questions (i.e., before and after the self-paced break), as described in Ben-Shakhar and Elaad (2002). Buffer and catch items were not included in the standardization procedure (see klein Selle et al., 2016, 2017). As the dependent measure, CIT-detection scores were calculated for each participant and each physiological measure by averaging the standardized score of all critical items. Since recognition of the critical item results in heart rate deceleration, HR CIT-scores are multiplied by -1 prior to analysis. Therefore, for both SCR and HR, a positive CIT-score is indicative of enhanced physiological responding upon the presentation of the critical item (i.e., concealed information recognition).

1.3.1. Exclusion criteria

On participant level, physiological data were eliminated from analyses when anomalies occurred (for HR, $n = 1$). For exclusions on trial level, for each of the dependent measures, item-specific responses were removed if the standardized score was smaller than -5 or larger than 5, reflecting outliers. When a movement coincided with a positive standardized score (for SCR) or a positive standardized score larger than 2 or lower than -2 (for HR), the item was discarded from analyses (see also Geven, klein Selle, Ben-Shakhar, Kindt, & Verschuere, 2018).

Further exclusions on response level were performed when participants showed a standard deviation of the raw SCR scores below 0.01 during the presentation of a block (i.e., 4 questions). In these cases, all SCR measurements from that block were discarded from further analyses due to non-responsiveness. Following these exclusion criteria, 96 % (range from 50 %–100 %) of the SCR and 97 % (range from 95 %–100 %) of the HR data were included in the analyses.

2. Results

All analyses used an alpha level of 0.05. Effect sizes for the independent samples t-tests are reported using Cohen's *d*. In addition, JZS Bayes factors (BF) were computed using JASP software version 0.8.4, representing numerical values quantifying the odds ratio between the null and the alternative hypothesis given the data. BF_{01} annotates how much more likely the data are under the null as compared to the alternative hypothesis, and BF_{10} annotates how much more likely the data are under the alternative as compared to the null hypothesis. For one-tailed testing, Bayes factors are reported as either predicting the null (BF_{0+}) or the alternative hypothesis (BF_{+0}). JZS prior with scaling factor $r = 1.000$ was used for the alternative hypothesis (see Rouder, Speckman, Sun, Morey, & Iverson, 2009). It should be noted that values close to 1 fail to support either hypotheses, and that Bayes factors represent a relative comparison of one hypothesis over the other but that both could be false.

2.1. Confirmatory analyses

2.1.1. True versus false confessions

It was expected that true confessors, but not false confessors, display physiological evidence of recognizing intimate details of the transgression in the CIT. We quantified the mean detection score such that positive CIT-scores provide physiological evidence of recognition. Therefore, it was expected that for both physiological measures the CIT-score to the critical items will be significantly higher for true confessors compared to false confessors. This was analyzed with a one-tailed

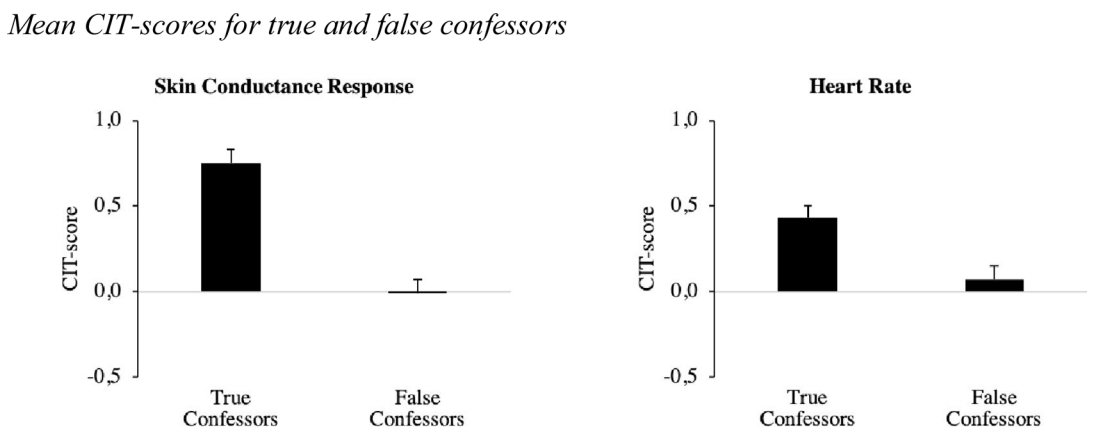

Fig. 3. Mean CIT-scores for true and false confessors.
Note. Error bars represent standard errors.

independent-samples *t*-test with Condition as the grouping variable and the CIT-score as the dependent measure. Fig. 3 shows the SCR and HR CIT-scores for false confessors and true confessors.

For the SCR, a significantly higher CIT-score was revealed for true confessors (M = 0.75, SD = 0.43) than for false confessors (M = -0.01, SD = 0.36), $t(46) = 6.41$, $p < .001$, $d = 1.88$, 95 % CI [1.18;2.56], BF_{+0} = 352863. For HR, true confessors (M = 0.43, SD = 0.38) showed a higher CIT-score than false confessors (M = 0.07, SD = 0.36), $t(45) = 3.35$, $p < .001$, $d = 1.00$, 95 % CI [0.37;1.61], BF_{+0} = 41.18.

Within each of the two conditions, we investigated whether participants exhibited significantly larger responses to the critical item compared to the control items (i.e., CIT-score > 0). For this purpose, a one-tailed one-sample *t*-test was conducted for true confessors and a two-tailed one-sample *t*-test for false confessors. We expected that a CIT-score significantly larger than zero will be observed for true confessors, reflecting recognition, whereas the response to the critical items should not differ from the control items amongst false confessors.

For SCR, true confessors indeed showed a mean CIT-score significantly larger than 0, $t(27) = 9.11$, $p < .001$, $d = 1.72$, 95 % CI [1.22;∞], BF_{+0} = 2.75e+7. For the false confessors the mean CIT-score did not significantly differ from zero, $t(19) = -0.15$, $p = .883$, d = -0.03, 95 % CI [-0.47;0.41]. Bayesian analysis showed evidence for the null hypothesis in false confessors, $BF_{01} = 5.80$.

For HR, true confessors showed a mean CIT-score significantly larger than 0, $t(27) = 6.08$, $p < .001$, $d = 1.15$, 95 % CI [0.74;∞], BF_{+0} = 23371. For the false confessors the mean CIT-score did not significantly differ from zero, $t(18) = 0.79$, $p = .442$, $d = 0.18$, 95 % CI [-0.28;0.63]. Bayesian analysis showed evidence for the null hypothesis in false confessors, $BF_{01} = 4.27$.

2.1.2. Individual detection accuracy

To analyze the detection efficiency of classifying participants as true or false confessors, we compared the distribution of the CIT-scores for both groups. Receiver operation characteristics (ROC) curves were constructed for each physiological measure separately, depicting the true positive rate versus the false positive rate for every possible classification threshold (see Supplementary Appendix C). As an outcome measure, we computed the area under the curve (*a)* using the mean CIT-score as the dependent variable and Condition as the state variable. This way, the accuracy of the CIT to classify true confessors and false confessors was calculated. The value of the area under the curve varies between 0 and 1, with a value of 0.5 representing classification at chance level and a value of 1 reflecting perfect accuracy in separating true from false confessors based on their physiological responding. For both physiological measures, the ROC is expected to be significantly above chance level, inferred when the lower boundary of the 95 % confidence interval is higher than chance (i.e., a value of 0.50).

Analyses revealed that detection efficiency was significantly larger than chance classification, with a = 0.90 [0.81;0.99] for SCR and a = 0.77 [0.63;0.90] for HR.

2.2. Exploratory analyses

2.2.1. Memory

All but one of the true confessors recalled both correct answers of the quiz (M = 1.96, SD = 0.19); none of the false confessors was able to do so (M = 0.00, SD = 0.00). No statistical analyses could therefore be conducted for recall. Among true confessors, 100 % correctly chose all trivia answers in a recognition task. In contrast, only four out of 20 false confessors choose the correct answer ‘Snapchat’ and three out of 20 false confessors correctly picked ‘Asia’. A one-tailed independent-samples *t*-test revealed that true confessors (M = 3.96, SD = 0.19) had a significantly higher recognition rate than false confessors (M = 0.65, SD = 1.14), $t(19.75) = 12.91$, $p < 0.001$, $d = 4.08$, 95 % CI [2.66;∞], BF_{+0} = 3.56e+16.

2.2.2. Motivation

Two-tailed independent-samples *t*-tests revealed that true confessors reported significantly more motivation to hide knowledge of the correct answers compared to false confessors. Additionally, true confessors reported significantly more effort to suppress and enhance physiological reactions during the test compared to false confessors. No significant differences were found between true and false confessors in the reported focus on the computer screen during the CIT, general involvement in the study, or memory for the answers of the trivia quiz. Table 1 shows the descriptive statistics.

2.2.3. Stress

Two-tailed independent-samples *t*-tests revealed that false confessors reported a significantly more negative experience than true confessors. Additionally, no significant differences were found between true and false confessors in the reported stress experienced during the interrogation. Two-tailed independent-samples *t*-tests revealed no significant differences between false confessors and true deniers regarding their experience and stress during the interrogation. Table 2 shows the descriptive statistics.

3. Discussion

In the current study, we examined the potential of the CIT to evaluate the veracity of confessions. The CIT uses physiological measures to verify whether the suspect recognizes critical details derived from the crime scene investigation, that would be known to the

Table 1
Descriptive statistics of the motivation questionnaire items (5-point Likert scale).

Question	*M (SD)*	*M (SD)*	*t*	*df*	*p*	$d_{between}$	*BF*
	True Confessors (n = 28)	False Confessors (n = 20)					
Effort to hide knowledge of the correct answers	4.21 (1.03)	2.50 (1.67)	4.07	29.23	< .001	1.24 [0.53;1.83]	BF_{10} = 316.33
Effort to suppress physiological responding	3.89 (0.99)	2.45 (1.28)	4.40	46	< .001	1.29 [0.65;1.91]	BF_{10} = 325.18
Effort to enhance physiological responding	2.64 (1.28)	1.70 (1.03)	2.72	46	.009	0.80 [0.20;1.39]	BF_{10} = 5.17
Focus on the computer screen	3.61 (0.96)	3.35 (0.99)	0.91	46	.370	0.27 [-0.31;0.84]	BF_{01} = 2.47
Involvement in the experiment	4.04 (0.74)	4.10 (0.79)	−0.29	46	.775	−0.08 [-0.66;0.49]	BF_{01} = 3.33
Memory for answers in the individual task	4.39 (0.74)	4.05 (0.61)	1.77	45.04	.084	0.51 [-0.07;1.10]	BF_{01} = 1.07

perpetrator. As a consequence, innocent suspects should not show physiological signs of recognition. The CIT produced an effect of d = 1.88 for the oldest and most often used measure in CIT research (SCR). This means that true confessors, but not false confessors, displayed physiological changes that are associated with intimate knowledge of the experimental transgression. The results indicate that, in the current study, there is a 91 % chance that a true confessor will have a higher SCR score than a person randomly picked from the false confessors.

The success of this laboratory study raises the question whether the CIT can and should be used in the forensic field to falsify confessions. A key challenge for the application of the CIT in real life investigative interviews is the fact that false confessions often contain information that is presumed to represent the perpetrator's knowledge of the crime (Garrett, 2010, 2015). Yet in a sample of 66 false confessions in DNA exoneration cases in the United States, the vast majority of false confessions contain details about the crime that are both accurate and not in the public domain (Garrett, 2015). In these instances, the presence of these details indicates that police had contaminated the process of interrogation by communicating information, providing secondhand guilty knowledge to otherwise innocent suspects (for an example, see Trainum, 2014).

It is critical to remember that the CIT assesses recognition of crime scene information, not guilt (Ogawa, Matsuda, & Verschuere, 2015). Results of several studies have revealed that the mere presence of knowledge suffices to elicit a CIT effect, irrespective of the source of the memory (e.g., Bradley, Barefoot, & Arsenault, 2011). In a mock-crime study, Bradley and Rettinger (1992) split participants into three groups in which the origin of critical information was manipulated. Participants who had committed a mock-crime could be clearly separated from innocents. Yet, those who were informed of the critical details without actually committing the act showed a larger CIT effect in comparison to innocents. While informed innocents could still be distinguished from guilty participants on a group level, individual classification on a case-to-case basis remains challenging. If no procedures are in place to prevent or observe the process of contamination in the interrogation room (see Alceste et al., 2020), such as recording police interrogations to document the original source of each detail appearing in the confession, innocents might be classified as knowledgeable in the CIT.

Under what conditions, then, may the CIT be still of use to evaluate the veracity of confessions? Only when investigators withhold detailed crime scene information during an investigation so that the suspect's recognition of intimate details as revealed by the CIT would have high evidential value (Osugi, 2011). In fact, the same conditions are necessary for applications of the CIT in forensic investigations. In case of partial contamination, there may remain non-contaminated details, or the CIT could consist of more detailed items (see also Osugi, 2018). Consider the case of Brendan Dassey, in which investigators asked leading and highly suggestive questions about what had happened to the victim's head (Ricciardi & Demos, 2015). After several incorrect guesses, the detective reveals to Dassey that the victim was shot. In a subsequent CIT, questions could still be asked about the number of shots, as well as the specific firearm used. Initial findings regarding this matter reveal promising results, showing that probing for more specific exemplar-level details (e.g., revolver) as compared to possibly contaminated broader information (e.g., firearm) might reduce the risk that contamination negatively affects CIT classification (Geven, Ben-Shakhar, Kindt, & Verschuere, 2019).

Moreover, using pictorial items enlarges the flexibility of the CIT (Lykken, 1998) and has been previously found suitable to detect concealed information (Ambach, Bursch, Stark, & Vaitl, 2010; Seymour & Kerlin, 2008). The presentation of pictorial items on an exemplar-level (e.g., images of various revolvers) may therefore provide a safeguard for innocents contaminated with categorical verbal information (e.g., that the suspect was shot). However, critical items may be less discernible from control items at an exemplar level than at a categorical level (Osugi, 2018) and further research has to specifically validate the use of visual exemplar-level details.

3.1. Limitations and future directions

This study is not without limitations. First, in the current paradigm, the transgression consisted of participants helping a confederate on a task that should have been solved individually. Questions can be raised about the generalizability of a CIT regarding a minor, provoked transgression compared to high-stake serious offenses. Moreover, it should be noted that the transgression in the current paradigm was a prosocial act, not an antisocial act, as it was the confederate who initiated the cheating.

Second, the cover story was mainly held up for the first three parts of the experiment; the personality questionnaires, problem-solving tasks, and –most importantly– the accusation and subsequent confession. For the final part of the experiment, all participants were informed that a deception detection test would be used to indicate whether they had cheated during the trivia quiz. Therefore, this information did not influence the decision to confess but was designed to ensure all

Table 2
Descriptive statistics of the stress questionnaire items (7-point Likert scale).

Question	*M (SD)*	*M (SD)*	*t*	*df*	*p*	$d_{between}$	*BF*
	True Confessors (n = 28)	False Confessors (n = 20)					
General experience of being accused of cheating	3.71 (1.12)	2.70 (0.92)	3.33	46	.002	0.97 [0.36;1.58]	BF_{10} = 19.71
Experienced stress while being accused of cheating	4.43 (1.64)	3.90 (1.71)	1.08	46	.286	0.32 [-0.26;0.89]	BF_{01} = 2.15
	True Deniers (n = 13)	False Confessors (n = 20)					
General experience of being accused of cheating	2.77 (1.36)	2.70 (0.92)	−0.17	31	.863	−0.06 [-0.76;0.64]	BF_{01} = 2.15
Experienced stress while being accused of cheating	3.31 (1.84)	3.90 (1.71)	0.94	31	.353	0.34 [-0.37;1.04]	BF_{01} = 2.11

participants responded with a verbal 'no' upon item presentation in the CIT. As guilty suspects are usually intrinsically motivated to conceal critical information, participants were explicitly instructed and incentivized to conceal their possible knowledge. For future research, participants could be instructed to verbally repeat all alternatives as this ensures active item processing without explicit deception.

Third, in the original paradigm by Russano et al. (2005) the confederate either actively sought assistance from participants, or not, resulting in a random assignment of guilt. In contrast, all participants in the current study were tempted by the confederate, thereby leading to self-selected rule-breaking behavior. Whereas random assignment to condition is an important feature to exclude confounding variables in research, this also leads to loss of ecological validity. Especially in the field of legal psychology, deception and rule-breaking behavior is typically self-selected. Crucially, self-initiated cheating does not undermine detection efficiency in the CIT compared to instructed cheating (Geven et al., 2018).

Fourth, while we tested a student sample and investigated situational factors used to obtain confessions, we did not consider the impact of dispositional characteristics that might render some suspects particularly vulnerable (e.g., intellectual disability; Appelbaum & Appelbaum, 1994; see also Clare & Gudjonsson, 1995; Gross, Jacoby, Matheson, Montgomery, & Patil, 2005; Kassin, Perillo, Appleby, & Kukucka, 2015). In the sample of 66 DNA exoneration cases investigated by Garrett (2015), more than a third of the confessions were obtained from juveniles; 22 exonerees had a diagnosed intellectual disability or mental illness. Previous research suggests that the CIT can be reliably used with children and juveniles (Visu-Petra, Jurje, Ciornei, & Visu-Petra, 2016) as well as in the general prison sample (Suchotzki, Kakavand, & Gamer, 2018). Besides improving special procedures for interviewing vulnerable suspects, the validity of the CIT with vulnerable suspects (e.g., cognitive impairments) should be further investigated.

4. Conclusions

An admission of guilt is considered the most probative evidence at trial that by itself increases the chance of a conviction (Leo & Ofshe, 1998). In light of the number of wrongful convictions in which false confessions were a contributing factor, it is of utmost importance to verify whether the confessor is factually guilty of the crime or gave a false incriminating statement. However, it can be a challenging task to reliably differentiate between true and false confessions in the criminal justice system. The current findings reveal initial evidence that the CIT may be able to distinguish true from false confessions. While further research is needed before the CIT could be implemented in law enforcement settings, the present findings demonstrate that confessing cheaters could be meaningfully distinguished from innocent false confessors by verifying the presence of crime-related knowledge is a promising starting point. Important restrictions with regards to the contamination of suspects in the field must be considered, but we point to possible future usage of the CIT under specific circumstances.

Acknowledgements

This research was funded by a grant, no. A2-8, from Stichting Fonds Catharine van Tussenbroek awarded to Linda Marjoleine Geven and also supported by a grant, no. 238/15, from the Israel Science Foundation awarded to Gershon Ben-Shakhar. We cordially thank Bert Molenkamp for programming the experiment, and Tessa Cornwall, Nynke van Gorkum, Emmeke Kooistra, Malouke Kuiper and Eva Scherlippens for their assistance during data collection.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.biopsycho.2020.107902.

References

Alceste, F., Jones, K. A., & Kassin, S. M. (2020). Facts only the perpetrator could have known? A study of contamination in mock crime interrogations. *Law and Human Behavior, 44*(2), 129–142. https://doi.org/10.1037/lhb0000367.

Ambach, W., Bursch, S., Stark, R., & Vaitl, D. (2010). A concealed information test with multimodal measurement. *International Journal of Psychophysiology, 75*(3), 258–267. https://doi.org/10.1016/j.ijpsycho.2009.12.007.

Appelbaum, K. L., & Appelbaum, P. S. (1994). Criminal justice-related competencies in defendants with mental retardation. *The Journal of Psychiatry & Law, 22*(4), 483–503. https://doi.org/10.1177/009318539402200402.

Appleby, S. C., Hasel, L. E., & Kassin, S. M. (2013). Police-induced confessions: An empirical analysis of their content and impact. *Psychology Crime and Law, 19*(2), 111–128. https://doi.org/10.1080/1068316X.2011.613389.

Ashton, M. C., & Lee, K. (2009). The HEXACO–60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*(4), 340–345. https://doi.org/10.1080/00223890902935878.

Ben-Shakhar, G., & Elaad, E. (2002). Effects of questions' repetition and variation on the efficiency of the guilty knowledge test: A reexamination. *The Journal of Applied Psychology, 87*(5), 972–977. https://doi.org/10.1037/0021-9010.87.5.972.

Bradley, M. T., & Rettinger, J. (1992). Awareness of crime-relevant information and the guilty knowledge test. *The Journal of Applied Psychology, 77*(1), 55–59. https://doi.org/10.1037/0021-9010.77.1.55.

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. https://doi.org/10.1017/CBO9780511975196.011.

Bruton v. United States, 391 U.S. 123 (1968).

Burns, S., McMahon, D., & Burns, K. (2012). *The Central Park Five [motion picture]*. US: Florentine Films.

Clare, I. C. H., & Gudjonsson, G. H. (1995). The vulnerability of suspects with intellectual disabilities during police interviews: A review and experimental study of decision--making. *Mental Handicap Research, 8*(2), 110–128. https://doi.org/10.1111/j.1468-3148.1995.tb00149.x.

Clow, K. A., & Leach, A. M. (2015). Stigma and wrongful conviction: All exonerees are not perceived equal. *Psychology Crime and Law, 21*(2), 172–185. https://doi.org/10.1080/1068316X.2014.951645.

Drizin, S. A., & Leo, R. A. (2004). The problem of false confessions in the post-DNA world. *North Carolina Law Review, 82*, 891–1007. https://doi.org/10.1525/sp.2007.54.1.23.

Exline, R. V., Thibaut, J., Hickey, C. B., & Gumpert, P. (1970). Visual interaction in relation to machiavellianism and an unethical act. In R. Christie, & F. L. Geis (Eds.). *Studies in machiavellianism* (pp. 53–76). New York, NY: Academic Press.

Garrett, B. L. (2010). The substance of false confessions. *Stanford Law Review, 62*, 1051–1119.

Garrett, B. L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong.* Cambridge, MA: Harvard University Presshttps://doi.org/10.4159/hardvard.9780674060982.

Garrett, B. L. (2015). Contaminated confessions revisited. *Virginia Law Review, 101*(2), 395–454.

Geven, L. M., Ben-Shakhar, G., Kindt, M., & Verschuere, B. (2019). It's a match!? Appropriate item selection in the Concealed Information Test. *Cognitive Research: Principles and Implications, 4*(1), 11. https://doi.org/10.1186/s41235-019-0161-8.

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. *Biological Psychology, 138*, 146–155. https://doi.org/10.1016/j.biopsycho.2018.09.005.

Gross, S. R., Jacoby, K., Matheson, D. J., Montgomery, N., & Patil, S. (2005). Exonerations in the United States 1989 through 2003. *The Journal of Criminal Law & Criminology, 95*(2), 523–560.

Grounds, A. (2005). Understanding the effects of wrongful imprisonment. *Crime and Justice, 32*, 1–58. https://doi.org/10.2307/3488358.

Gudjonsson, G. H. (1989). Compliance in an interrogative situation: A new scale. *Personality and Individual Differences, 10*(5), 535–540. https://doi.org/10.1016/0191-8869(89)90035-4.

Gudjonsson, G. H. (2003). Psychology brings justice: The science of forensic psychology. *Criminal Behaviour and Mental Health, 13*(3), 159–167. https://doi.org/10.1002/cbm.539.

Hasel, L. E., & Kassin, S. M. (2009). On the presumption of evidentiary independence: Can confessions corrupt eyewitness identifications? *Psychological Science, 20*(1), 122–126. https://doi.org/10.1111/j.1467-9280.2008.02262.x.

Honts, C. R., Forrest, K., & Stepanescu, A. (2019). Polygraph examiners unable to discriminate true and false juvenile confessions. *Polygraph and Forensic Credibility Assessment: A Journal of Science and Field Practice, 48*(1), 1–9.

Honts, C. R., Kassin, S. M., & Craig, R. A. (2014). "I'd know a false confession if I saw one": A constructive replication with juveniles. *Psychology Crime and Law, 20*(7), 695–704. https://doi.org/10.1080/1068316X.2013.854792.

Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2013). *Essentials of the reid technique.* Burlington, MA: Jones & Bartlett Publishers.

Kassin, S. M. (2015). The social psychology of false confessions. *Social Issues and Policy Review, 9*(1), 25–51.

Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law and Human Behavior, 21*(1), 27–46. https://doi.org/10.1023/A:1024814009769.

Kassin, S. M., & Wrightsman, L. S. (1980). Prior confessions and mock juror verdicts. *Journal of Applied Social Psychology, 10*(2), 133–146. https://doi.org/10.1111/j.1559-1816.1980.tb00698.x.

Kassin, S. M., & Wrightsman, L. S. (1981). Coerced confessions, judicial instruction, and mock juror verdicts. *Journal of Applied Social Psychology, 11*(6), 489–506. https://doi.org/10.1111/j.1559-1816.1981.tb00838.x.

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition, 2*(1), 42–52. https://doi.org/10.1016/j.jarmac.2013.01.001.

Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law and Human Behavior, 29*(2), 211–227. https://doi.org/10.1007/s10979-005-2416-9.

Kassin, S. M., Perillo, J. T., Appleby, S. C., & Kukucka, J. (2015). *Confessions. APA handbook of forensic psychology, Vol. 2: Criminal investigation, adjudication, and sentencing outcomes*245–270. https://doi.org/10.1037/14462-009.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, 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.

klein Selle, N., Verschuere, B., Kindt, M., Meijer, E., & Ben-Shakhar, G. (2017). Unraveling the roles of orienting and inhibition in the Concealed Information Test. *Psychophysiology, 54*(4), 628–639. https://doi.org/10.1111/psyp.12825.

Leo, R. A., & Ofshe, R. J. (1998). The consequences of false confessions: Deprivations of liberty and miscarriages of justice in the age of psychological interrogation. *The Journal of Criminal Law & Criminology, 88*(2), 429. https://doi.org/10.2307/1144288.

Lykken, D. T. (1959). The GSR in the detection of guilt. *The Journal of Applied Psychology, 43*(6), 385–388. https://doi.org/10.1037/h0046060.

Lykken, D. T. (1974). Psychology and the lie detector industry. *The American Psychologist, 29*(10), 725–739. https://doi.org/10.1037/h0037441.

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

Marion, S. B., Kukucka, J., Collins, C., Kassin, S. M., & Burke, T. M. (2016). Lost proof of innocence: The impact of confessions on alibi witnesses. *Law and Human Behavior, 40*(1), 65–71. https://doi.org/10.1037/lhb0000156.

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. https://doi.org/10.1111/psyp.12239.

Neuschatz, J. S., Lawson, D. S., Swanner, J. K., Meissner, C. A., & Neuschatz, J. S. (2008). The effects of accomplice witnesses and jailhouse informants on jury decision making. *Law and Human Behavior, 32*(2), 137–149. https://doi.org/10.1007/s10979-007-9100-1.

Ogawa, T., Matsuda, I., & Verschuere, B. (2015). The Concealed Information Test in the laboratory versus Japanese field practice: Bridging the scientist-practitioner gap. *Archives of Forensic Psychology, 1*(2), 16–27. https://doi.org/10.16927/afp.2015.1.2.

Osugi, A. (2011). Daily application of the concealed information test: Japan. In B. Verschuere, E. Meijer, & G. Ben-Shakhar (Eds.). *Memory detection: Theory and application of the concealed information test* (pp. 253–275). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9780511975196.015.

Osugi, A. (2018). Field findings from the concealed information test in Japan. In J. P. Rosenfeld (Ed.). *Detecting concealed information and deception: Recent developments* (pp. 97–121). London, UK: Academic Press. https://doi.org/10.1016/B978-0-12-812729-2.00005-7.

Ricciardi, L, & Demos, M (Producers) (2015). Making a murderer. US: Netflix.

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*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225.

Russano, M. B., Meissner, C. A., Narchet, F. M., & Kassin, S. M. (2005). Investigating true and false confessions within a novel experimental paradigm. *Psychological Science, 16*(6), 481–486. https://doi.org/10.1111/j.0956-7976.2005.01560.x.

Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S. D., & Castiello, U. (2008). How to accurately assess autobiographical events. *Psychological Science, 19*(8), 772–780. https://doi.org/10.1111/j.1467-9280.2008.02156.x.

Scott, L. (2010). "It never, ever ends": The psychological impact of wrongful conviction. *American University Criminal Law Brief, 5*(2), 10–22.

Seymour, T. L., & Kerlin, J. R. (2008). Successful detection of verbal and visual concealed knowledge using an RT-based paradigm. *Applied Cognitive Psychology, 22*, 475–490. https://doi.org/10.1002/acp.1375.

Sokolov, E. N. (1963). Perception and the conditioned reflex. *The American Journal of Psychology, 79*(1), 166. https://doi.org/10.2307/1420734.

Suchotzki, K., Kakavand, A., & Gamer, M. (2018). Validity of the reaction time concealed information test in a prison sample. *Frontiers in Psychiatry, 9*, 745. https://doi.org/10.3389/FPSYT.2018.00745.

Trainum, J. L. (2014). *How the police generate false confessions: An inside look at the interrogation room.* New York, NY: Rowman & Littlefield.

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

Visu-Petra, L., Jurje, O., Ciornei, O., & Visu-Petra, G. (2016). Can you keep a secret? Introducing the RT-based Concealed Information Test to children. *Psychology Crime and Law, 22*(3), 276–301. https://doi.org/10.1080/1068316X.2015.1109085.

Wetmore, S. A., Neuschatz, J. S., & Gronlund, S. D. (2014). On the power of secondary confession evidence. *Psychology Crime and Law, 20*(4), 339–357. https://doi.org/10.1080/1068316X.2013.777963.