UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Institut für Systemische Neurowissenschaften

Direktor: Prof. Christian Büchel

Emotional modulation of memory encoding and retrieval in the Concealed Information Test

Dissertation

zur Erlangung des Grades eines Dr. rer. biol. hum.

im Rahmen des PhD Programm für Nichtmediziner an der Medizinischen Fakultät der

Universität Hamburg.

vorgelegt von:

Judith Peth

aus Berlin

Hamburg, Januar 2014

(wird von der Medizinischen Fakultät ausgefüllt)

Angenommen von der

Medizinischen Fakultät der Universität Hamburg am: 23.6.2014

Veröffentlicht mit Genehmigung der

Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Christoph Mulert

Prüfungsausschuss, zweite/r Gutachter/in: Dr. Matthias Gamer

Prüfungsausschuss, dritte/r Gutachter/in: Prof. Dr. Jan Wacker

Table of Contents

List of Abbreviations

ANS	autonomic nervous system
CIT	Concealed Information Test
CQT	Control Question Test
CTP	Complex Trial Protocol
ERP	event-related potential
fMRI	functional magnetic resonance imaging
FPWL	finger pulse waveform length
FWHM	full width at half maximum
IFG	inferior frontal gyrus
ISI	inter stimulus interval
MFG	middle frontal gyrus
NRC	national research council
OR	orienting response
PFC	prefrontal cortex
RLL	respiration line length
ROI	region of interest
RT	reaction time
SCR	skin conductance response
SMG	supramarginal gyrus
TPJ	temporal parietal junction

List of Tables

List of Figures

1. Introduction

The polygraph is a device that continuously measures several physiological responses to investigate whether a persons is deceiving or tells the truth (for a historical review see Ben-Shakhar & Furedy, 1990, p.1 ff.). In 2003, the US National Research Council (NRC) claimed, that “The theoretical rationale for the polygraph is quite weak, especially in terms of differential fear, arousal, or other emotional states that are triggered in response to relevant or comparison questions.” (p. 213). In other words, the NRC is concerned that the scientific evidence for usage of the polygraph is not appropriate to enable conclusions for realistic settings outside of the laboratory. Indeed, huge amounts of research focus on the question, to what extend physiological responses enable a differentiation between somebody who tells the truth and somebody who lies. Based on inter-individual variability in physiological arousal, no single physiological measure enabled such a differentiation. Therefore, question techniques became an important instrument, because they are thought to permit a differentiation between responses on different question types. Especially the “Control Question Test” (CQT; also referred to as the Comparison Question Test) and the “Concealed Information Test” (CIT; also referred to as Guilty Knowledge Test) are influential approaches in this line of research.

Most of the concerns raised by the US NRC (2003) pertain to CQT studies. The CQT is the most commonly used polygraph method in several countries (e.g., the United States of America and Israel). This approach compares physiological responses (e.g., respiration, skin conductance, relative blood pressure) on two different question types. The relevant question asks if the suspect committed the crime (e.g., “Did you break into Mr. Jones house last Friday night?”). In contrast, the control question asks for a probable lie (e.g., ”Have you ever taken something that did not belong to you?”). To deny this question should induce emotional stress in the person and is used as a comparison for the relevant question. The basic idea of the CQT is that guilty subjects show stronger responses for relevant compared to control questions. For innocent persons the opposite response pattern is predicted. In addition to these two question types, irrelevant questions are presented that are not analyzed (e.g., “Are you sitting on a chair?”). More detailed descriptions of the CQT, the respective data analysis and interpretation can be found elsewhere (e.g., Lykken, 1998; Raskin, 1989; Reid & Inbau, 1966). The CQT's validity to correctly classify guilty subjects (i.e., sensitivity) differs between laboratory and field studies, as well as the correct classification of innocent persons (i.e., specificity). For example, laboratory studies reported sensitivity scores ranging from 50% to 92% (weighted average: 80%, n = 238) and specificity scores ranging from 39% to 85% (weighted average: 63%, n = 249). In contrast, 76% to 94% sensitivity (weighted average: 84%, n = 319) and 20% to 91% specificity (weighted average: 72%, n = 282) were found for scientific field studies (cf., Ben-Shakhar and Furedy, 1990). Strikingly, the successful outcome of the CQT depends on the investigator's ability to increase a person's uncertainty while denying the control questions. Aim of the CQT is an increased stress level in the suspect. Thus, this technique is highly sensitive for characteristics of both, the investigator as well as the person under investigation. Moreover, innocent subjects have a high risk to show stronger responses to relevant questions, for example based on feelings of threat during the investigation. Therefore, false positive rates above 40% were reported for the CQT (Patrick & Iacono, 1991).

In contrast, the Concealed Information Test (CIT) mainly relies on recognition of crime-related information and not on the manipulation of emotional states in a suspect. The CIT consists of a series of multiple-choice questions, which ask for specific details of a crime under investigation (e.g., “Which entry did you use to break into Mr. Jones house last Friday night?”) that should only be known to a guilty person and the police/investigator. The single answer options should be equally plausible to an innocent (i.e., not involved) examinee (Lykken, 1959, 1998). For the given example, answer options could be: a) the main door, b) the kitchen window, c) the cellar door, d) the balcony, e) the roof window. Only a guilty subject is supposed to recognize the critical crime-related detail and respond with a specific physiological reaction, consisting of skin conductance increase, respiratory suppression, and heart rate deceleration (Gamer, Rill, Vossel, & Gödert, 2006). Laboratory research revealed valid detection of concealed information with the CIT and reported 82% sensitivity and 93% specificity under optimized study conditions (cf., Ben-Shakhar & Elaad, 2003). However, there is a lack of studies addressing aspects of realistic crime scenarios. For example, it remains rather unclear to what extend the recognition of details is influenced by increased emotional arousal during commitment of a crime or a longer delay between crime and CIT investigation. Additionally, the encoding of critical information during a crime could be influenced by situational characteristics. For example, it is unclear whether an emotional context or real enactment compared to intention affects the depth of information processing. Aside from studies that explore the CIT under rather realistic conditions, results from basic research on emotional modulation of memory can be suitable to develop hypotheses for the emotional modulation of crime-related memory. Therefore, the current thesis aimed at linking applied CIT research with basic research on emotional memory. The first study investigated the influence of emotional arousal on the encoding of relevant details during a mock-crime and the physiological responses for these details during a CIT. In addition, some subjects participated in a delayed CIT to reflect real-life conditions. The second study compared responses during the CIT between guilty subjects and a group of informed innocents, after both groups partly encoded the same details in different contexts (i.e., criminal or non-criminal). Additionally, a guilty intention group was investigated to compare responses for relevant details that were either encoded during planning or real enactment. The two studies used different physiological and behavioral response measures (e.g., autonomic response measures, neuroimaging, eye-tracking) as dependent measure during the CIT.

1.1 The Concealed Information Test (CIT)

The CIT was developed by David Lykken (1959) and aims at detecting crime-related memory while minimizing biases due to confounding factors (e.g., investigators effects). Usually, the CIT includes five to ten questions. Every question consists of one relevant detail and four to five irrelevant answer options. Different details of a specific crime might be suitable for construction of CIT questions, for example the masking of an offender, the car used to escape from a crime scene, or the amount of stolen money. Usually, the subject is asked to respond with “no” to each single answer option while data of different autonomic response measures are collected during the investigation. This includes the measurement of respiration via belts placed around the chest and/or the stomach, peripheral blood flow recording in the fingers, attachment of electrodes to the chest to record an electrocardiogram and attachment of electrodes to the palm of the hand to measure electrodermal responses. After the subject recognizes the relevant detail, a decrease in heart rate, respiration and peripheral blood flow is observed, accompanied by increased electrodermal responses (Gamer et al., 2006). Although each single measure enables a valid detection of concealed information, a combined score of electrodermal, respiratory and heart rate measures reached better results compared to single scores (Gamer, Verschuere, Crombez, & Vossel, 2008). Over the last decades, several other measures were found to detect concealed memory. Among these are reaction times (RTs; e.g., Seymour, Seifert, Shafto, & Mosmann, 2000; Verschuere, Crombez, Degrootte, & Rosseel, 2010), event related potentials (ERPs; e.g., Farwell & Donchin, 1991; Seymour et al., 2000) or functional magnetic resonance imaging (fMRI; e.g., Gamer, Klimecki, Bauermann, Stoeter, & Vossel, 2012; Nose, Murai, & Taira, 2009). Most of these results base on laboratory studies, and there is a lack of research on real-life applications of the CIT.

1.1.1 Scientific background of the CIT

A correctly constructed CIT enables a differentiation between relevant and neutral answer options that matches the definition of comparison cues in a scientific view, because uninformed persons should not be able to differ between these two item types (cf., Ben-Shakhar & Elaad, 2003). Accordingly, a group of naïve persons can be used to check whether the single answer options are indistinguishable for innocent subjects. In general, the CIT is highly accepted in the scientific community, because it is based on solid scientific principles and numerous publications support its empirical basis (Iacono & Lykken, 1997).

The theoretical rational behind the CIT is the orienting response (OR; Sokolov, 1963) which includes behavioral and physiological responses, elicited by rare, novel or personally significant stimuli or by changes in stimulation. Sokolov (1963) hypothesized that a person builds an internal representation of a repeatedly presented stimulus. Input that mismatches the characteristics of this representation results in an OR. In contrast, input that matches the representation inhibits the OR and habituation takes place. In the CIT, relevant details of a crime become salient for a guilty person. Thus, the presentation of these stimuli among irrelevant details should evoke enhanced ORs (Lynn, 1966; Sokolov, 1963). In general, the OR theory explains autonomic response changes in guilty subjects after presentation of relevant details of a crime during the CIT very well. However, it does not explain all observed physiological changes. For example, the heart rate deceleration after presentation of relevant details during the CIT is temporally much more extended than predicted by OR theory and some expected physiological changes (e.g., increased startle responses for relevant details) were not supported by empirical evidence (Verschuere, Crombez, Koster, Van Bockstaele, & De Clercq, 2007). Therefore, alternative concepts like response inhibition might be useful to explain specific physiological changes after presentation of relevant details during the CIT (cf., Verschuere & Ben-Shakhar, 2011). Response inhibition is defined as an executive function that allows a person to intentionally inhibit a dominant, automatic or prepotent response (Miyake et al., 2000). Evidence from RT-based CIT studies supports the idea that guilty subjects have to inhibit their initial response during information concealment (e.g., Noordraven & Verschuere, 2013; Verschuere & De Houwer, 2011). Existing literature therefore suggests that responses in the CIT are not only driven by ORs, instead higher cognitive functions are assumed to additionally explain detection of guilty knowledge in the CIT.

Regarding the validity of the CIT in the laboratory, two meta-analyses on electrodermal responses were published. MacLaren (2001) collected data from 22 studies and reports a sensitivity of 76% (*n* = 843) and a specificity of 83% (*n* = 404 for the CIT). Most of the reported studies used a simple analysis measure, the Lykken scoring (Lykken, 1959). For this measure, the single answer options of a question are assessed with 2 points for the strongest response, 1 point for the second strongest response and 0 points for all additional responses. MacLaren (2001) included flexible cut-off scores, defined by the respective study, to classify whether the response scores belong to the guilty or innocent category. Thus, comparison between studies is limited. A second meta-analysis by Ben-Shakhar and Elaad (2003), included data from 80 laboratory studies (*N* = 5198) and is therefore the most extensive work on the CIT's validity to date. In contrast to prior analyses, effect sizes were used and pooled over studies to enable statistical comparisons between different approaches, independent from the cut-off criteria employed by the respective studies. Most of the included CIT studies investigated memory for artificial stimuli, such as autobiographical events, prior learned word lists or playing cards. In addition, the authors collected a sub-sample of data based on mock-crime paradigms, in which subjects are asked to commit an instructed mock-crime. As this study design was found to closest resemble the real-life situation, it is considered the ideal approach to investigate detection of concealed information. In addition, this meta-analysis revealed some conditions that increase the validity of the CIT (cf., Ben-Shakhar & Elaad, 2003). Thus, the test should include at least five questions on relevant details, subjects should be instructed to verbally respond “no” to each presented detail instead of staying silent, and motivation to beat the CIT should be increased by offering rewards for successful participation. For a sub-set of studies fulfilling these criteria, Ben-Shakhar and Elaad (2003) reported an average effect size of d = 3.12 (compared to d = 1.55 over all included studies). Cohen (1988) defined as rule of thumb, that a threshold of 0.20 reflects small effects, 0.50 medium effects and 0.80 high effects. Therefore, the CIT is a highly effective procedure to detect concealed information. However, these results are limited to electrodermal response measures. More recent research showed that inclusion of additional autonomic measures, like respiration and heart rate changes, increased the respective effect sizes (e.g., Gamer, 2011a; Gamer et al., 2008).

As mentioned before, laboratory research strongly supports the validity of the CIT, but there is a lack of field studies investigating the validity of the CIT under less optimized conditions (for a review see Elaad, 2011a). So far, two field studies examined the CIT hit rates in real-life settings (Elaad, 1990; Elaad, Ginton, & Jungman, 1992). These studies reported sensitivity scores of 50% and 76% and specificity scores of 98% and 94%, respectively. Importantly, these results were collected in police interrogations that did not constitute optimal CIT conditions. For example, the CIT occurred after investigation with a CQT. In average only two relevant details were used, which did not match the recommended amount of at least five details (cf., Ben-Shakhar & Elaad, 2003). Thus, it is still questionable how good the CIT performs in real-life settings when optimal conditions as identified by Ben-Shakhar and Elaad (2003) are taken into account.

1.1.2 CIT construction in the field

Japan is the only country that implemented the CIT in their regular police routine. Here, 5000 CIT investigations are conducted per year (Osugi, 2011). The CIT is a suitable method to decrease the amount of suspects for a certain crime and test results might work as legal argument or can proof somebody’s innocence. The practical construction of a CIT can be divided into five steps, where the first two steps focus on test preparation and do not include any contact to a person under investigation. The following section describes the single steps of construction in further detail and uses examples from Japan (e.g., Osugi, 2011; Nakayama, 2002) for illustration purposes.

Step 1: Extraction of relevant crime details

To construct proper CIT questions, the investigator has to gather as much information as possible about the crime and specific details of the scene. Therefore, the investigator visits the crime place, talks to eye witnesses or crime victims and should use every source to get relevant information (e.g., video tapes from security cameras). This way, the investigator collects details that a guilty subject should remember. At least five relevant details are necessary to guarantee valid detection of concealed information (cf., Ben-Shakhar & Elaad, 2003). Depending on the type of crime, different details might be suitable for the CIT construction (cf., Nakayama, 2002). Questions on a murder investigation could include the crime place, the murder method, the place where the corpse was found or the way the corpse was hidden. Since the choice of relevant details is one of the most important aspects during CIT construction, some criteria for relevant details were formulated by Osugi (2011): (1) Details cannot be known from a different context than the crime (e.g., public media; interrogation during arrest); (2) Details are central aspects of the crime that are easy to remember; (3) Details are suitable to validate the legal issue of the crime and should be directly connected to the crime under investigation (e.g., the stolen object).

Step 2: Construction of the CIT

To develop multiple-choice question sets for the CIT, the instructor has to choose equally plausible neutral answer options for each critical detail. For an innocent person, these answer options should be indistinguishable from the respective relevant detail. Stimulus presentation can occur auditory or visual, or as a combination of both modalities. The items should be standardized and resemble each other regarding specific picture characteristics (e.g., brightness, usage of salient colors, sound volume). The position of the relevant details among the irrelevant details should be counterbalanced to prevent anticipation effects. Most importantly, an irrelevant detail should always be in the first position after question presentation, as usually a stronger physiological response occurs after presentation of the question. Usually, the first answer option serves as a buffer item and is not included in the data analysis. It is advisable to construct as many question sets as possible to enable the instructor to leave questions out or change questions if necessary (cf., Lykken, 1998, p. 288 ff.). To prevent leakage of relevant details, no critical details of a crime under investigation should be given to the public media or be mentioned in suspect interviews prior to the CIT.

Step 3: Pre-Test Interview

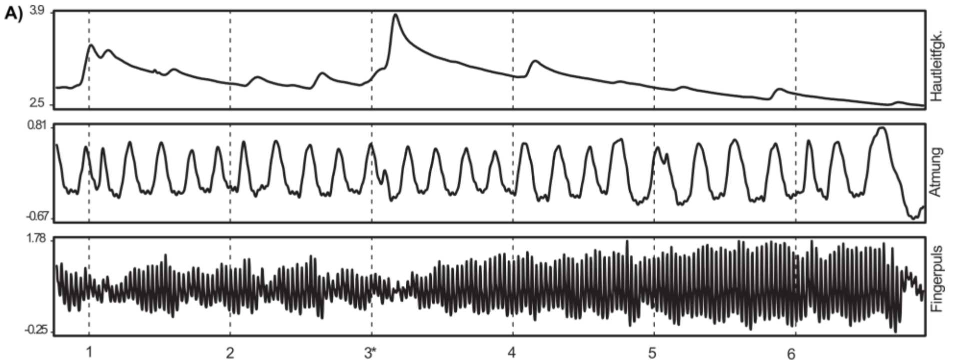
Before the actual CIT investigation, a pre-test interview is conducted. Here, autobiographic information of the suspect and aspects that might influence the responses during the CIT session (e.g., prior drug consumption, amount of hours sleep) are documented. This interview should ensure the suspect's ability to participate in the CIT and written informed consent to participate in the CIT has to be signed. In addition, the investigator explains the measurement procedure and tries to reduce potential fears or tensions regarding the CIT. The content of the CIT questions is presented to the suspect and the person can tell the investigator whether he or she got knowledge about relevant details from media reports or other sources. If these explanations are plausible, the respective questions can be discarded from the CIT and be replaced with other questions.

Step 4: CIT investigation

The investigator attaches electrodes and other necessary equipment to the suspect's body and presents the prior defined question sets during the CIT. The delay between single answer options is about 20 to 25 s. The presented visual stimuli can be photographs or real objects (Nakayama, 2002). Autonomic responses (e.g., respiration, peripheral blood flow, cardiovascular responses, electrodermal responses) are measured continuously during the test. Modern polygraph systems enable digital data collection and storage while the older systems document the response changes on paper. Each question set is presented three to five times with a typical CIT session duration of 60 to 90 minutes, depending on the amount of questions and repetitions. The investigator should be aware of countermeasures a suspect might use. Countermeasures are deliberate techniques that are used by a suspect to alter the physiological reactions during the CIT. Therefore, each physical movement (e.g., toes and fingers) or artificial sound (e.g., clearing the throat) during answering should be documented and the suspect should be instructed to stop such behavior (Osugi, 2011).

Step 5: Data analysis and post-test interview

According to the theoretical background of the CIT, only persons who were involved in the crime should be able to distinguish between relevant and irrelevant answer options. Figure 1A and 1B depict typical responses in guilty persons. As shown in Figure 1C and 1D, innocent subjects cannot distinguish between item types and an unsystematic response pattern is expected for these persons (Lykken, 1998). Thus, systematic responses are interpreted as recognition of relevant information and the respective analysis can occur with automated computer programs. After the CIT, a post-test interview should be conducted to clarify whether some answer options had a specific meaning for the suspect. After a reasonable explanation, the respective details can be excluded from further analysis.



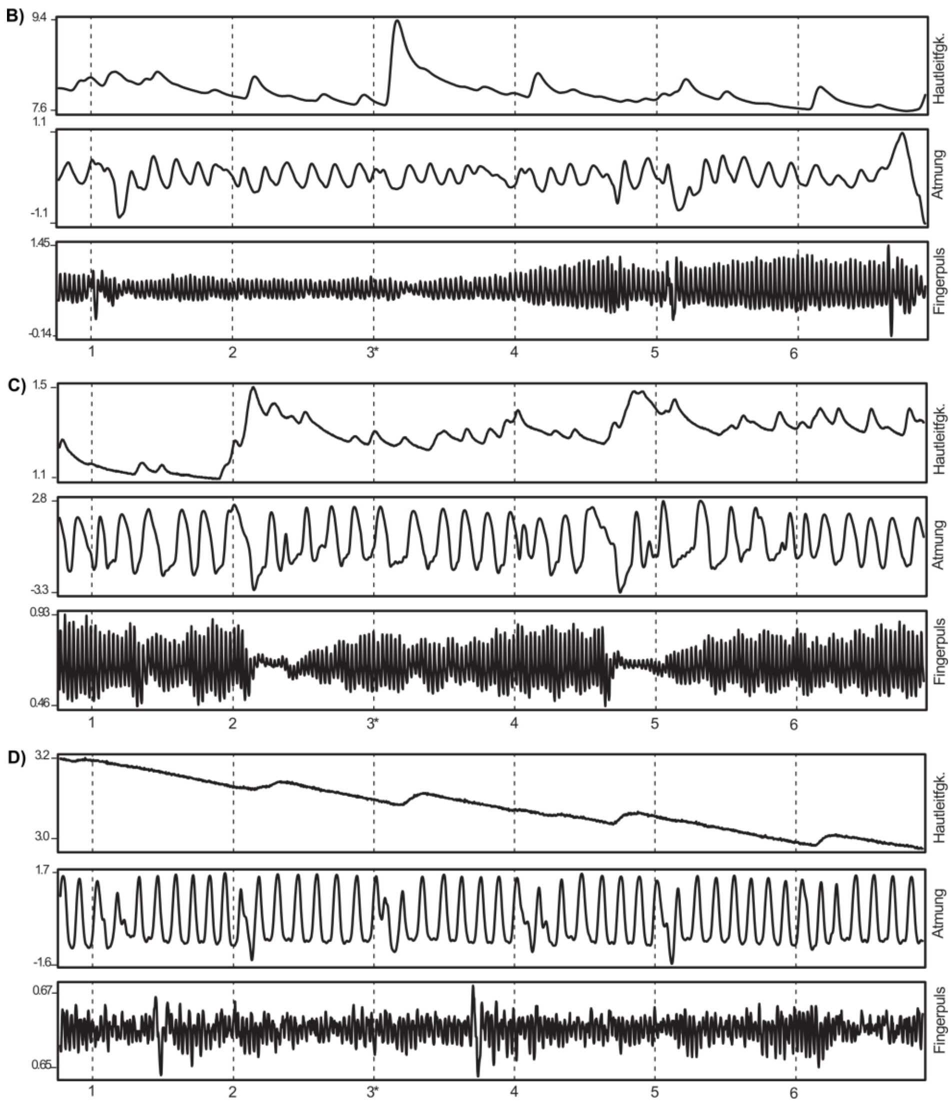


Figure 1. Raw data responses of guilty and innocent persons (investigated in Study I). Starting with the upper line, each picture contains of the skin conductance, the respiration and the finger pulse waveform. The question was: “What did you steal from the room?” and the answer options were: 1. Projector, 2. Mobile phone, 3*. Money, 4. External hard disk, 5. Camera, 6. Laptop. Option 3* was the relevant detail. Dotted lines indicate the presentation of single answer options. A) This guilty subject shows a strong decrease in finger pulse volumes after presentation of the correct answer option 3*, accompanied by an increase in skin conductance. Afterwards a general decline in reactivity occurs. B) In addition to the increase in skin conductance and the decrease in finger pulse volumes, this guilty subject shows a suppression of the respiratory activation after presentation of answer 3*. C) This innocent subject shows an unsystematic response pattern, including stronger responses for some answer options. D) This innocent subject shows a constantly pronounced decline in skin conductance, without specific responses for single answer options. Adapted from “Aktuelle Forschung zur Validität des Tatwissentests: Der Einfluss von Emotionen” by J. Peth and M. Gamer, 2013, *Praxis der Rechtspsychologie*, 23 (1), p. 151-165. Copyright by Deutscher Psychologen Verlag GmbH. Reprinted and adapted with permission.

1.2 Physiological and behavioral response measures for the CIT

The CIT originally relied on measures of the autonomic nervous system to reveal concealed knowledge and these measures are the most frequently used approach until today. Over the last decades, additional measures were found to enable the detection of concealed information with a CIT. For example, in line with the increased amount of research on neural processes, fMRI-based (e.g., Gamer et al., 2012; Nose et al., 2009) and ERP-based (e.g., Farwell & Donchin, 1991; Rosenfeld et al., 1988) CIT studies were conducted. In addition, explicit behavioral approaches like reaction time measures (e.g., Seymour et al., 2000; Verschuere et al., 2010) were investigated, as well as covert measures that are not perceived by the person under investigation (e.g., Elaad & Ben-Shakhar, 2008, 2009).

1.2.1 Autonomic response measures

The autonomic nervous system (ANS) is responsible for synchronization of peripheral functions (Öhman, Hamm, & Hugdahl, 2000). It consists of sympathetic and parasympathetic branches, which are generally associated with activation and relaxation, respectively. The ANS is sensitive for emotional factors, but includes also rather general functions like digestion, homeostasis, effort and attention (cf., Berntson & Cacioppo, 2000).

In his first study, Lykken (1959) used only skin conductance response (SCR) changes to detect crime-related information. He correctly classified 100% of the innocent subjects and 88% of the guilty subjects. The second group committed a mock-crime before the test. Since then, most CIT studies used electrodermal responses and correspondingly the two meta-analysis on the CIT (MacLaren, 2001; Ben-Shakhar & Elaad, 2003) mainly included SCR studies. SCR changes are strongly correlated with activity in the sympathetic nervous system (Wallin, 1981). Thus, the sympathetic nervous system in guilty subjects seems to be stronger activated during presentation of crime-related information compared to neutral information. Lykken (1974) assumed, that an OR (Sokolov, 1963) causes these specific reactions, because the relevant details of a crime reach high salience for a guilty suspect. This hypothesis was supported by several electrodermal response characteristics present during the CIT. Among these were the habituation over time and the recovery after presentation of another stimulus (Barry, 1996), as well as increased skin conductance amplitudes for significant stimuli (Siddle, O'Gorman, & Wood, 1979).

In addition to electrodermal responses, other measures of the autonomic nervous system were examined as dependent variables during the CIT. Timm (1982) found that respiration activity could be used to detect concealed information by calculating the respiration line length (RLL) based on the length of the respiration tracing for 10 to 15 s after stimulus presentation. This measure decreases when breathing gets slower or when the respiratory amplitude is reduced. CIT studies using this response measure demonstrated that the RLL in guilty subjects shows a stronger decrease after presentation of crime relevant details compared to irrelevant details (cf., Gamer, 2011a). Interestingly, respiratory measures were reported to be more sensitive to emotional factors compared to electrodermal responses and therefore the field validity of the RLL might be higher (Elaad et al., 1992). Furthermore, the RLL was reported to be more resistant against the usage of countermeasures during the CIT (cf., Ben-Shakhar, 2011). Respiration is mainly regulated by parasympathetic structures that interact with central structures and peripheral feedback circuits (Lorig, 2007). Additionally, respiratiory suppression was reported to show some characteristics related to the OR, like occurrence after presentation of unexpected stimuli (cf., Gamer, 2011a).

Other autonomic measures frequently used in the CIT are cardiovascular ones. Especially changes in phasic heart and pulse rate were found to be valid detectors of concealed knowledge, as both measures decreased systematically in guilty persons after presentation of relevant details (e.g., Ambach, Stark, Peper, & Vaitl, 2008; Bradley & Ainsworth, 1984; Bradley & Janisse, 1981; Verschuere, Crombez, De Clercq, & Koster, 2004). Interestingly, the heart rate was reported to show a biphasic response pattern after presentation of relevant details during a CIT. The initial heart rate increase is followed by a heart rate deceleration. Only the later part of the heart rate response validly detected concealed information and is thought to reflect OR (Verschuere, Crombez, Smolders, & Clercq, 2009). Another cardiovascular measure, the finger pulse, was repeatedly found to enable differentiation between persons (e.g., Gamer, Verschuere, et al., 2008; Podlesny & Raskin, 1977). Measurement occurs with a photoplethysmograph, which is attached to one finger of the suspect's hand. Different measures can be calculated from these data, for example the pulse rate, the finger pulse amplitude or the finger pulse waveform length (FPWL). Some CIT studies demonstrated a systematically smaller FPWL in guilty subjects after presentation of relevant details compared to irrelevant details (e.g., Elaad & Ben-Shakhar, 2006). However, FPWL validity estimates are lower compared to electrodermal estimates (e.g., Ambach et al., 2008; Elaad & Ben-Shakhar, 2006, 2008; Verschuere et al., 2009). The autonomic regulation of cardiovascular responses involves the sympathetic and the parasympathetic nervous system (cf., Berntson, Quigley, & Lozano, 2007). While heart rate changes are controlled by both branches of the autonomic nervous system, peripheral vasoconstricton is mainly under control of the sympathetic nervous system (cf., Gamer, 2011a).

In general, the OR theory explains the presented changes in autonomic responses during a CIT very well, but recent research questions this association as the only explanation for physiological changes during the test (e.g., Gamer, Gödert, Keth, Rill, & Vossel, 2008; Verschuere, Crombez, Koster, Van Bockstaele, et al., 2007).

A promising approach to increase the validity of the CIT is the combination of multiple autonomic measures. Heterogeneous findings were reported on the optimal combination of autonomic measures (e.g., Elaad, 2009; Verschuere, Crombez, Koster, & De Clercq, 2007). However, Gamer and colleagues (2008) used a stepwise logistic regression approach to identify the optimal weight for the most popular autonomic measures during a CIT. They found that the ideal regression coefficients should be set to β = -3.917 for the constant, $\beta = 4.24$ for electrodermal responses, $\beta = -6.310$ for respiration, and $\beta = -1.975$ for the heart rate. This model was cross-validated later on, supporting the finding that the combined measure outperforms the single autonomic measures (Gamer, Verschuere, et al., 2008). From a scientific view, it is not fully clear why the combined score reaches better results than the best single measure. One explanation could be that a combination of these autonomic measures covers different aspects of information concealment and enables a better adaptation to individual differences in physiological responsiveness between subjects (cf., Gamer, 2011a).

1.2.2 Event-related potentials

Electroencephalographic recordings of stimulus-dependent changes in brain waves, the so-called ERPs, enable an examination of neurocognitive processes within a few hundred milliseconds after stimulus presentation (Bressler & Ding, 2006). For the detection of concealed information, the P300 component became very important (for a review see Rosenfeld, 2011). This ERP component is characterized by a large positivity in the brain wave signal 300 to 500 ms after stimulus onset and is associated with stimulus salience (Pritchard, 1981).

The traditional CIT compares responses to relevant and irrelevant details that are presented only a few times to prevent habituation effects. The inter stimulus interval (ISI) is rather long and lasts from 20 to 25 s (cf., Ben-Shakhar & Elaad, 2003). P300-based CITs (e.g., Allen, Iacono, & Danielson, 1992; Donchin & Coles, 1988; Farwell & Donchin, 1991; Rosenfeld et al., 1988) require a transformation of the traditional test design to comply with the requirements of an ERP study. More precisely, an increased number of stimulus repetitions and smaller ISIs are necessary. In addition, target items are presented that require a different button press compared to all other items (i.e., relevant and irrelevant details). The targets are learned by the subjects prior to the CIT and should increase the subject's motivation and attention during the test. All items are presented repeatedly on the screen and the subjects are instructed to press the respective button (e.g., left button for targets, right button for all other items) as quickly as possible. However, this frequently used three-stimulus-paradigm was reported to have some limitations, especially the usage of countermeasures decreased the validity of this approach (cf., Rosenfeld, 2011).

To overcome this problem, the *complex trial protocol* (CTP; Rosenfeld et al., 2008) was developed and is reported to be unaffected by countermeasure usage (Labkovsky & Rosenfeld, 2012; Rosenfeld & Labkovsky, 2010). The CTP consists of two sequential tasks, of which the first is the critical one. Usually, the subject is asked to respond immediately to a presented stimulus by pressing one out of two buttons (e.g., press the left one). The respective stimulus is either a relevant detail or an irrelevant detail, but subjects respond equally to them. This first response is instructed to the subjects as pure perception confirmation. Approximately 1 s after presentation of the first stimulus, another stimulus is presented. This stimulus is either a target, that requires a different response (e.g., press the right button) or a non-target item that requires the same button press as the first stimulus (e.g., press the left button). The subject has to respond as quickly as possible. Important for the CTP, the first stimulus response should reveal a stronger P300 for relevant compared to irrelevant details in guilty subjects. The second stimulus response only maintains attention and ensures task compliance.

Moreover, recent studies investigated other ERP components in CIT settings. For example, Gamer and Berti (2010) used a three-stimulus-protocol in a playing card CIT. They found increased N200 amplitudes for remembered relevant details compared to irrelevant details, which enabled the valid detection of concealed knowledge. This finding was replicated with auditory stimulus presentation during the CIT (Matsuda, Nittono, Hirota, Ogawa, & Takasawa, 2009). However, the N200 effect was reported to depend on a diminished distinctiveness between the presented relevant and irrelevant details (Gamer & Berti, 2012). Thus, the N200 component is interpreted as an indicator for increased cognitive control, including conflict monitoring (Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003), but was not assumed to mirror processing of relevant crime details in general (Gamer & Berti, 2012). This finding further supports the assumption that different cognitive mechanisms underlie the CIT and the OR theory is not able to explain all phenomena visible in the CIT. Furthermore, central and autonomic nervous system components might detect slightly different aspects of the same phenomena. For example, differences in sensitivity for the depth of information processing in a mock-crime scenario were reported for electrodermal responses, but not for ERP measures (Gamer and Berti, 2012). Thus, a combination of different measures could increase the detection of concealed information by covering different psychological aspects that are measured during the CIT.

Overall, the P300-based CIT is a valid procedure to detect concealed information (cf., Rosenfeld, 2011). The external validity of this approach is still unclear, because the majority of P300 studies focused on autobiographical details and less accurate detection rates were reported for the ERP-based CIT in mock-crime designs (Rosenfeld, Biroschak, & Furedy, 2006; Rosenfeld, Shue, & Singer, 2007). Whereas accuracy rates were between 85% and 95% in the laboratory (Rosenfeld, Soskins, Bosh, & Ryan, 2004), the only field study reported an accuracy rate of approximately chance level (Miyake, Mizutani, & Yamahura, 1993).

1.2.3 Behavioral measures

In contrast to autonomic response measures, behavioral measures are usually expected to strongly depend on a person’s conscious control. However, these measures offer some advantages (e.g., low cost for data collection; simple measurement without application of electrodes to the subject’s body) in contrast to the traditional approaches. For example, reaction time (RT) measurement is a behavioral approach to examine a person’s response to a stimulus and was reported to enable the detection of concealed information under specific circumstances (for a review see Verschuere & De Houwer, 2011). The RT-based CIT design usually resembles the design developed for the ERP-based CIT. Therefore, the comparison between response latencies for relevant and irrelevant details is critical for the detection of concealed information, but target details are included to increase the subject's attention during the test by requiring a different response button press. Only guilty subjects were reported to have problems in categorizing relevant details as non-targets, represented by slower response times and increased error rates (e.g., Seymour et al., 2000; Seymour & Fraynt, 2009; Seymour & Kerlin, 2008). For the RT-based CIT a sensitivity from 81% to 90% and a specificity from 85% to 98% were reported (cf., Verschuere and De Houewer, 2011). These results are comparable to electrodermal detection rates in mock-crime CIT investigations. Interestingly, the difference between target items and irrelevant items could be used to investigate whether subjects are classifiable by a CIT (Noordraven & Verschuere, 2013). In sum, the RT-based CIT can be a valid detector of concealed information, but this approach has some limitations. For example, the RT performance strongly depends on the subject's motivation to respond as fast and correct as possible. Since such compliance is not expected in guilty subjects, this measure might be less applicable in field settings (cf., Matsuda et al., 2012).

In contrast, implicit measures like eye-tracking devices often enable measurement of behavioral reactions without attachment of any equipment to the person's body. So far, some studies have shown that the pupillary response (e.g., Bradley & Janisse, 1981; Lubow & Fein, 1996) or the overall blinking behavior (e.g., Fukuda, 2001; Leal & Vrij, 2010) are sensitive to concealed knowledge. For example, Leal and Vrij (2010) examined blink activity while their participants made either true or false statements about the commitment of a mock-crime. They found, that liars showed fewer blinks for relevant questions compared to persons telling the truth. A discriminant analysis revealed a sensitivity of 75% and a specificity of 77% for this approach (Leal & Vrij, 2010). Fukuda (2001) measured the number of blinks participants produced on each trial during a CIT and the results showed that relevant details led to a higher average blink-rate compared to irrelevant details. Unfortunately, a detailed classification analysis was not reported, making it difficult to assess the validity of this approach. Pupil dilatation was used in a CIT as well, and enabled correct classification in 50% of the guilty and 100% of the innocent subjects (Lubow and Fein, 1996). This result shows that pupil measures might be interesting for usage in a CIT, but the accuracy rates are below rates from more established CIT measures. In sum, ocular measures seem to be promising approaches, but further research is needed to clarify to what extend they enable valid detection of memory during a CIT. To investigate this issue, the first study of the current thesis included eye-movements and eye-blink recordings in addition to the autonomic response measures during the CIT.

1.2.4 Functional brain-imaging

Neuroimaging techniques became very important for investigating the neural mechanisms underlying memory detection (for a review see Gamer, 2011b). In contrast to the ERP's, which have a high temporal resolution, functional magnetic resonance imaging (fMRI) has a high spatial resolution. Therefore, anatomical structures relevant for the detection of concealed information can be investigated in further detail. It is an open question whether direct measures of neural activity could outperform autonomic measures.

Most of the CIT studies using fMRI employed a card-test design (e.g., Gamer, Bauermann, Stoeter, & Vossel, 2007; Langleben et al., 2002; Nose et al., 2009), where subjects had to conceal knowledge of a specific playing card that was presented among different other cards. In most of these studies, rarely presented target items required a different behavioral response and were included to ensure the participant's attention during stimulus presentation. This design closely resembles the ERP-based CIT design (e.g., Farwell & Donchin, 1991; Rosenfeld et al., 1988). Across different fMRI studies, relevant items elicited higher activity then neutral alternatives in a ventral fronto-parietal brain network, consisting of the bilateral inferior frontal gyrus (IFG), the right middle frontal gyrus (rMFG) and the right temporoparietal junction (rTPJ) (cf., Gamer, 2011b). It is important to note that these regions are not exclusively involved in deception or the concealment of knowledge. For example, the IFG is assumed to reflect detection of relevant stimuli in a stream of irrelevant ones (Kiehl, Laurens, Duty, Forster, & Liddle, 2001) as well as retrieval of relevant details from memory (Iidaka, Matsumoto, Nogawa, Yamamoto, & Sadato, 2006). Activity in the TPJ was reported to reflect changes in the environment (Downar, Crawley, Mikulis, & Davis, 2000, 2002). Interestingly, the ventral fronto-parietal network was found to be activated while presenting relevant items during a CIT independent of response demands (Gamer et al., 2012). This delivers further support for the assumption that neural responses during a CIT are mainly based on memory processes and occur independent of motivational aspects.

Most fMRI studies employing the CIT reported group comparisons and so far, only two studies investigated individual classification accuracy using fMRI responses during a CIT. Nose and colleagues (2009) reported 84% sensitivity as well as 84% specificity based on activity differences between relevant and neutral alternatives in the right ventrolateral prefrontal cortex (PFC). Ganis and colleagues (2011) showed that the right lateral PFC and the anterior medial PFC could be used to correctly classify 100% of the subjects, guilty and innocent. Importantly, these accuracy rates were dramatically reduced for guilty subjects that were trained to use physical countermeasures during the CIT (e.g., move the index finger of the left hand). Of these guilty subjects, only 33% were classified correctly (Ganis, Rosenfeld, Meixner, Kievit, & Schendan, 2011).

In sum, the fMRI-based CIT revealed that the difference in neural activation between relevant irrelevant details is mainly driven by memory-retrieval processes (Gamer et al., 2012). Nevertheless, there is a lack of fMRI research investigating the influence of emotional, motivational and cognitive factors on a suspect's performance during the CIT. For example, no mock-crime study using fMRI was published so far. Since this paradigm was shown to approximate realistic settings best (Ben-Shakhar & Elaad, 2003), such a study is necessary to further estimate the external validity of the fMRI-based CIT. The second study of the current thesis aimed to investigate these issues. Therefore, neural activity during the CIT was measured in persons that encoded the relevant details during specific contexts (i.e., criminal versus non-criminal; enactment versus intention).

1.2.5 Covert measures

Some studies investigated the potential of covert measures (i.e., secret recording of a person's physiological responses) to detect concealed information. This approach might be very important for real-life applications of the CIT, because it does not require attachment of sensors to the suspect's body and a decreased usage of countermeasures by the suspects is assumed. Because covert measures enable examinations without prior information of the investigated person, the person under investigation is assumed to further concentrate on verbal responses or facial expressions instead of autonomic responses (for a review see Elaad, 2011b).

Elaad and Ben-Shakhar (2008) delivered evidence for the successful usage of covert respiration measures during a CIT. They placed two hidden respiratory sensors in the seat and in the back support of a chair and calculated the RLL for subjects who were investigated with a CIT while sitting on this chair. This response measure enabled a valid detection of concealed knowledge. In a similar study, the covert respiratory measures during the CIT were reported to be resistant to mental countermeasures but not physical countermeasures (Elaad & Ben-Shakhar, 2009).

Another covert measure reported to reveal guilty knowledge is facial temperature (Pollina et al., 2006). However, the validity of this measure was questioned based on the high drop-out rate and the obtrusive character of the study design (cf., Elaad, 2011b). As mentioned in section 1.2.3 of the current thesis, recording of ocular measures can be covert measure that enables a valid detection of concealed information (e.g., Fukuda, 2001; Leal and Vrij, 2010).

Overall, further evidence regarding the usage of covert measures during real-life CIT settings is needed to further investigate which measures are useful for concealed information detection. However, in this context concerns regarding ethical and legal issues become very important, for example a suspect usually gives written informed consent before participating in a CIT. Therefore, the measurement of physiological responses without the suspect's awareness could become ethically problematic.

1.3. Memory and the CIT

Explicit memory of the relevant details of a crime is important to ensure a valid detection of concealed information with the CIT. Where concerns regarding emotional and motivational factors stronger apply to approaches like the CQT, memory is particularly relevant for the CIT. For a long time, most laboratory CIT studies ensured that participants remembered relevant details by repeated presentation of these details prior to the test (cf., Ben-Shakhar & Elaad, 2003). Additionally, most CIT studies applied the test immediately after the subject learned the relevant information. Compared to realistic settings, these conditions are rather artificial. For example, suspects in real crimes would encode salient, relevant details only during commitment of the crime and the CIT would probably occur weeks or months later. Thus, there is a lack of studies examining how relevant details are perceived and encoded in real-life settings and whether situational factors (i.e., increased emotional arousal during the commitment of the crime) modulate the physiological response pattern in the CIT.

1.3.1 The CIT in real-life settings

Only a few CIT studies tried to set up experimental conditions in the laboratory that approximate the field situation to a certain degree (Carmel, Dayan, Naveh, Raveh, & Ben-Shakhar, 2003; Gamer, Kosiol, & Vossel, 2010; Nahari & Ben-Shakhar, 2010). In these studies, most (Carmel et al., 2003) or all relevant details (Gamer et al., 2010; Nahari & Ben-Shakhar, 2011) were incidentally encoded during commitment of a mock-crime and some participants underwent a CIT examination one or two weeks later to approximate field conditions. All studies reported a better detection of guilty participants when using CIT questions asking for central details of the crime compared to peripheral ones. Central details are directly related to the mock-crime, while peripheral details are present on the scene but are not directly involved in commitment of the mock-crime. Heterogeneous results were reported regarding the comparison between immediate and delayed CIT examinations. Whereas one study found a decline in validity coefficients as a function of time when the mock-crime was accomplished under realistic conditions (Carmel et al., 2003), the other studies could only replicate this result for a CIT based on peripheral information (Gamer et al., 2010; Nahari & Ben-Shakhar, 2011). However, when confining the test to central crime details, CIT validity was temporally stable in these latter studies.

Regarding the long-term detection of concealed information with the CIT, Hu and colleagues (2012) reported no decline in detection efficiency for one critical detail (i.e., the stolen good) in a P300-based CIT one month after the mock-crime (Hu, Hegeman, Landry, & Rosenfeld, 2012). A report from Japan demonstrated temporal stability of a P300-based CIT when the test was delayed up to one year (Hira, 2003). This study also asked only for the stolen good. Thus, it is currently unknown whether other aspects of a crime could also be reliably detected in a CIT examination after such long periods of time.

1.3.2 The emotional modulation of memory

Basic research repeatedly reported an advantage in memory for emotional details compared to neutral details. This emotional enhanced memory effect had been replicated for different stimuli, like words, pictures and narrated slide shows (for a review see Hamann, 2001). Interestingly, this benefit in memory was more pronounced for central details and is accompanied by an expense of memory for more peripheral details (cf., Christianson, 1992). In general, an increase in overt attention towards central aspects during an emotional context was assumed to explain this phenomenon together with an attentional shift away from irrelevant aspects (e.g., Christianson, 1992; Kensinger, 2009; Reisberg & Heuer, 2004), but empirical support for this hypothesis was rather weak (e.g., Riggs, McQuiggan, Farb, Anderson, & Ryan, 2011; Steinmetz & Kensinger, 2013). However, Kim and colleagues

(2013) reported that central details were better remembered compared to peripheral details, even when they were processed with less attention. This effect was especially pronounced in an emotional context and the authors conclude that the encoding of an emotional context seem to qualitatively differ from the encoding of a neutral context (Kim, Vossel & Gamer, 2013).

At a neural systems level, the prefrontal areas, the hippocampus and the amygdala were found to have specific meaning for the encoding and retrieval of emotional memories (for a review see Dolcos, Denkova, & Dolcos, 2012). Activity in limbic brain areas like the amygdala is reported to correlate with later memory for these stimuli (e.g., Cahill et al., 1996; Phelps & LeDoux, 2005) and emotional arousal was found to mediate the amygdala and hippocampus activation during memory encoding (Dolcos, LaBar, & Cabeza, 2004). Emotionally arousing stimuli were reported to result in a more stable memory trace based on increased hippocampal consolidation (Kensinger, 2009), as proven by longer retention delays for such details compared to neutral details (LaBar & Phelps, 1998). However, less conclusive neuroimaging data is available regarding emotional modulation of retrieval processes (for a review see Buchanan, 2007), but activation in amygdala, hippocampus and PFC were reported to reflect successful emotional memory (Dolcos, Denkova, & Dolcos, 2012). Importantly, the emotionally enhanced memory effect was found to require a consolidation-delay between encoding and retrieval (Ritchey, Dolcos, & Cabeza, 2008; Sharot, Verfaellie, & Yonelinas, 2007). Overall, central aspects of an event were found to be better remembered in emotional settings, while peripheral aspects were forgotten more often (Adolphs, Denburg, & Tranel, 2001; Phelps & Sharot, 2008).

In sum, emotion and memory seem to interact during all stages of information processing and result in the recruitment of specific brain regions as well as differences in explicit memory. Since the majority of studies on the emotional enhancement memory effect used stimuli like pictures or sentences (cf., Kensinger, 2009; Hamann, 2001), it remains unclear to what extent these results are transferable to a realistic CIT setting. Because the commitment of a crime can be understood as an emotional event, implications for the processing of relevant details of a crime and their detection during the CIT are assumed.

1.4 Research questions

The role of emotional modulation in encoding of relevant details of a crime and their retrieval during the CIT is still unclear. Therefore, the current thesis includes two studies that aimed to shed light on these questions. The first study investigated the validity of the CIT under realistic conditions, including manipulations of emotional arousal during the mock-crime and of the amount of time until application of the CIT. In addition to autonomic measures, eye-movements were used as an unobtrusive approach to measure concealed information. The second study aimed to investigate the association between context and depth of information processing on memory for relevant details during the CIT. For this purpose an fMRI-based CIT approach was used to investigate the influence of encoding context under real-life conditions. The following sections explain the mentioned research questions in more detail.

1.4.1 Influence of emotional arousal on the CIT (Study I)

Although laboratory studies found high validity coefficients for the CIT in differentiating between innocent and guilty subjects (Ben-Shakhar & Elaad, 2003; MacLaren, 2001) its external validity is still debated (Honts, 2004). The main point of criticism emphasized that laboratory studies were usually carried out under conditions that optimized the participants' recognition of relevant crime details during the CIT examination. For example, prior studies guaranteed recognition of relevant details and conducted the CIT immediately after the mock-crime (cf., Ben-Shakhar and Furedy, 1990, p.55-56). Thus, there is a lack of studies examining how relevant details are perceived and encoded during a criminal act and whether certain aspects of crimes in the field modulate the physiological response pattern in the CIT. Another aspect that has not yet been investigated by CIT studies is the role of emotional arousal for memory encoding. Although it can be expected that perpetrators experience enhanced arousal during a crime, it is currently unknown to what extent such arousal influences memory for crime-related information. In general, emotional arousal enhances memory (Kensinger, 2009). More specifically, it was found that emotional arousal strengthens long-term declarative memory for gist at the expense of memory about details (Adolphs, Denburg, & Tranel, 2001; Phelps & Sharot, 2008) and these effects require a certain delay, because they are modulated by memory consolidation (Hamann, 2001). The reported findings are in line with research on eyewitness memory, suggesting that increased arousal due to the presence of threatening objects in a situation might narrow the focus of attention to such central details (e.g., a weapon), which subsequently results in reduced memory of more peripheral information (Steblay, 1992). The first study of the current thesis aimed to examine to what degree these results can be generalized to a perpetrator's memory for crime details and his or her physiological responding in a CIT. For this purpose, four groups of guilty subjects were built based on manipulations of their emotional arousal during the mock-crime as well as the amount of time delay until the CIT was conducted. Half of the relevant details reflected central information of the mock-crime, while the second half reflected peripheral information. Differences in explicit memory for central and peripheral details were hypothesized, as well as an influence on autonomic response measures after presentation of these relevant details during the CIT.

1.4.2 Influence of encoding context on the CIT (Study II)

The second study of the current thesis aimed to investigate the relationship between encoding context and the respective detection of concealed information with an fMRI-based CIT. So far, only few studies investigated this issue, but none of them used fMRI. Some studies focused on the detection of criminal intentions, where encoding is assumed to be less deep compared to accomplished criminal actions. It was shown that skin conductance responses (Meijer, Verschuere, & Merckelbach, 2010), P300 amplitudes of ERPs (Meixner & Rosenfeld, 2011), and reaction times (Noordraven & Verschuere, 2013) revealed intentions to commit a mock-crime. Additionally, basic research on memory reported increased recognition for actively produced information (e.g., De Winstanley & Bjork, 1997; Slamecka & Graf, 1978) and a superior memory for actions over memory for verbally learned material (Engelkamp, 1998).

Another line of research focused on the potential of the CIT to differentiate between subjects who encoded relevant details in neutral situations or during criminal actions. This was usually done by comparing guilty subjects with informed innocents, who encoded the relevant details without committing a mock-crime (e.g., Ben-Shakhar, Gronau, & Elaad, 1999; Bradley & Rettinger, 1992; Bradley & Warfield, 1984; Gamer, 2010). In most of these studies, the relevant information was presented to the innocent subjects in a criminal context, for example by reading about the mock-crime or witnessing the crime. A meta-analytic review by Ben-Shakhar and Elaad (2003) reported false positive rates between 25% and 50% for informed innocents, which is dramatically higher than the 5% rate usually found (Ben-Shakhar & Elaad, 2003). More recently, Gamer (2010) reported indistinguishable autonomic responses of guilty subjects and informed innocents when crime-related knowledge was deeply encoded and participants were motivated to pass the test. The informed innocents witnessed the mock-crime by viewing a short video before they underwent the CIT (Gamer, 2010). Thus, even though informed innocents were not directly involved in the criminal act, the critical information was presented to them in the criminal context. By contrast, the relevant details of a crime could also be presented in a completely neutral context. This way, the informed innocents would realize during the CIT that they know details that belong to a crime. Giesen and Rollison (1980) as well as Stern, Breen, Watanabe, and Perry (1981) used this approach and reported a correct classification of 95% and 96% for guilty subjects and 100% and 88.5% for informed innocent subjects, respectively. Thus, the context of encoding seemed to have an influence on the CIT performance. However, other studies could not replicate these high hit rates and reported substantially lower validity estimates for informed innocents (e.g., Gamer, Gödert, et al., 2008). Overall, it is relatively unclear to which degree encoding of relevant details in a neutral context influences the physiological response pattern in the CIT.

To further investigate these questions, two groups of guilty subjects were examined in the second study of the current thesis. One group encoded relevant details only during the planning phase for a mock-crime which they unexpected did not enact. A second group fulfilled the mock-crime after an identical planning phase. In addition, a group of informed innocents was investigated. These persons encoded half of the relevant details during the planning and the enactment of an errand.

1.4.3 Application of eye-movements and eye-blinks during the CIT (Study I)

So far, only a few studies investigated eye tracking devices as a potential measure for CIT applications. This is surprising, since basic research demonstrated that eye movement patterns reflected memory across different stimulus classes, such as faces, buildings and scenes, even under circumstances in which behavioral reports may not (or cannot) be reliably obtained (for a review see Hannula et al., 2010). These studies reported a decrease in sampling behavior (e.g., fewer fixations, fewer regions viewed, increased fixation durations) for familiar as compared to unfamiliar stimuli (e.g., Althoff & Cohen, 1999; Ryan, Hannula, & Cohen, 2007).

With respect to CIT applications, only few reports indicated that eye-blinks and eye movement patterns might differ between relevant and neutral items when the examinee is able to identify the crime-related detail. Specifically, eye-blinks were suppressed for relevant as compared to neutral CIT items (Leal & Vrij, 2010) and a rebound effect was observed, with higher blink rates for relevant details after stimulus offset (Fukuda, 2001). With respect to other ocular measures, it was reported that the number of fixations tended to be reduced when participants viewed details that were associated to a previously accomplished mock-crime (Twyman, Moffitt, Burgoon, & Marchak, 2010). This effect, however, was comparably small and only occurred for one out of five relevant details. In a somewhat different experimental design, Schwedes and Wentura (2012) presented a set of six faces simultaneously on the screen and participants had to select one of them. Importantly, they were instructed to conceal knowledge of a specific familiar face that was also present on the display. It turned out that these familiar faces were fixated longer than the simultaneously presented distractors. This recognition effect was evident after the second fixation and the fixation pattern allowed for identifying the concealed faces in 65% of the trials while correctly classifying 92% of the trials that only contained unknown faces. To further investigate whether eye-movements and eye-blinks enable a valid detection of concealed knowledge, the first study included ocular measures (i.e., the number of fixations, the average fixation duration, the number of eye-blinks) in addition to the traditional autonomic measures.

2. Study I

2.1 Introduction

The first study of the current thesis investigated the validity of the CIT under rather realistic conditions. Laboratory studies found high validity coefficients for the CIT in differentiating between innocent and guilty subjects (Ben-Shakhar & Elaad, 2003; MacLaren, 2001). However, its external validity is still debated (Honts, 2004), because laboratory studies were usually carried out under conditions that optimized participants' recognition of relevant crime details during the CIT examination. Thus, there is a lack of studies examining how relevant details are perceived and encoded during a criminal act and whether certain aspects of crimes in the field modulate the physiological response pattern in the CIT.

As mentioned before, only a few laboratory studies tried to set up experimental conditions that approximate the field situation to a certain degree by assuring the incidental encoding of the relevant details (Carmel et al., 2003; Gamer et al., 2010; Nahari & Ben-Shakhar, 2011) during the commitment of a mock-crime. In the same studies, some participants underwent a CIT examination one to two weeks later to better resemble field conditions where the CIT is also conducted some time after the event. All studies reported a better detection of guilty participants when using CIT questions asking for central details of the crime (e.g., the amount of money stolen) compared to peripheral ones (e.g., a picture on the wall). Heterogeneous results were reported for the comparison between immediate and delayed CIT examinations. Whereas one study found a decline in validity coefficients as a function of time when the mock-crime was accomplished under realistic conditions (Carmel et al., 2003), the other studies could only replicate this result for a CIT based on peripheral information (Gamer et al., 2010; Nahari & Ben-Shakhar, 2011). When confining the test to central crime details, however, CIT validity was temporally stable in these latter studies.

Winograd and Rosenfeld (2011) delivered further support for CIT applications under realistic conditions by reporting positive results for a P300-based CIT targeting one salient item that was incidentally encoded during a mock-crime (Winograd & Rosenfeld, 2011). Overall, more realistic mock-crime studies involving incidental encoding of crime-related information demonstrated the advantage of memory for central over peripheral details and emphasized the importance of information encoding for the CIT outcome.

In the current study, participants incidentally encoded central and peripheral details while carrying out a realistic mock-crime. Emotional arousal during the mock-crime was manipulated in a between-subjects design and monitored using heart rate recordings and subjective nervousness ratings. Moreover, participants were tested either immediately after the mock-crime or with a delay of two weeks. For the arousal induction group as well as the control group, better memory and stronger physiological responses in the CIT for central as compared to peripheral details were expected. This difference should be more pronounced after additional induction of arousal during the mock-crime. Although previous studies yielded inconsistent results with respect to a delayed application of the CIT, it seems possible that enhanced arousal during the mock-crime strengthens long term memory for central aspects of the crime and increases corresponding physiological responses by modulating memory consolidation (Hamann, 2001).

A second aim of the current study was to clarify to what extent fixations and eye-blinks allow for a valid detection of concealed crime-related memories. So far, there are only few studies available that used ocular measures in a CIT examination, and these studies used heterogeneous experimental designs and provided partly inconsistent results. In order to clarify whether eye tracking data allows for a valid identification of crime-related memory in a standard CIT setting, this measure was investigated in addition to autonomic measures. Based on previous literature, a reduced blink rate (e.g., Leal & Vrij, 2010), a decrease in the number of fixations and longer fixation durations (e.g., Ryan, Hannula, & Cohen, 2007) were expected for pictures of relevant crime details compared to irrelevant details. Effects of the time and emotional arousal manipulations were investigated as exploratory analysis, but no previous hypotheses were generated regarding effects of these experimental manipulations.

2.2 Methods

2.2.1 Participants

This study was approved by the local ethics committee and conducted according to the principles expressed in the Declaration of Helsinki. All participants gave written informed consent and were paid for participation. Eighty-five subjects (25 women, 60 men) with a mean age of 26.3 years (*SD* = 3.6 years) participated voluntarily in the study. Most of them were students of different fields (76%). All participants had normal or corrected-to-normal vision.

2.2.2 Instruments

Physiological responses were recorded with a Biopac MP100 device (Biopac Systems, Inc.). Skin conductance was measured at the thenar and hypothenar eminences of the participant’s non-dominant hand by a constant voltage system (0.5 V) using a bipolar recording with two Hellige Ag/AgCl electrodes (surface area = 1cm^2) filled with 0.05 M NaCl electrolyte. Respiration was recorded with a piezo-electric belt attached around the participant’s chest. Finger pulse was measured with a photoplethysmograph, attached at the thumb of the non-dominant hand. In addition, an electrocardiogram was recorded using 3M RedDot Ag/AgCl electrodes attached to the manubrium sterni and the left lower rib cage. The reference electrode was placed at the right lower rib cage.

Eye movements during the CIT were monitored using a video-based eye tracker (EyeLink 1000, SR Research, Ontario, Canada) with a spatial resolution of less than 0.01° and a spatial accuracy of 0.5°. The eye tracking camera was placed under the display screen in front of the participant and the distance between camera and participant was tracked using a small target sticker affixed to the examinee’s forehead. This setup allowed for an effective sampling rate of eye movements with 500 Hz. The Software Presentation (Neurobehavioral Systems) was used to present the stimuli on a 19” LCD monitor, with a screen resolution of 1600 × 1200 pixels. Distance between participants and the screen was approximately 70 cm.

The physiological measurements were conducted in a sound-attenuated room. All recording and programming equipment was located outside the room and the participant's verbal responses could be monitored via a microphone installed in the room.

2.2.3 Design

A realistic mock-crime procedure was used in a 2 × 2 × 2 design with the between-subject factors arousal manipulation (arousal induction vs. no arousal induction) and delay between mock-crime and CIT (immediately vs. two weeks delayed) with 17 participants in each group. The type of critical detail (central vs. peripheral) was varied as a within-subject factor. For validity calculations, an additional group of 17 innocent examinees was tested after carrying out an alternative instruction that kept them ignorant of the relevant details. These subjects were all tested immediately. Participants were randomly assigned to experimental conditions by means of a predefined list.

2.2.4 Procedure

The experiment was conducted in two stages: stage 1 was constructed like a role play for the participants during which they had to fulfill a clearly defined task, stage 2 was the CIT application where they had to try to be classified as innocent.

Stage 1. Participants arrived at the laboratory individually, were met by an assistant who obtained written informed consent and were asked to rate their current subjective nervousness on a Likert scale ranging from 1 (*not nervous at all*) to 6 (*very nervous*). This measure was used as baseline. Subsequently, the assistant explained the experiment and assured that all participants understood the instruction for their particular condition. All participants wore a portable heart rate monitor (Polar RS800CX) in the form of a wristwatch that allowed for the precise measurement of the heart rate from beat to beat. The monitor included the option to set a marker during recording by pressing a button. Baseline heart rate was recorded for 120 s while participants were sitting still and relaxed. Before participants left the examination room they had to rate their current subjective nervousness again on the same Likert scale as before.

Guilty participants were instructed to commit a mock-crime in a small storeroom of the Department of Systems Neuroscience (University Medical Center Hamburg-Eppendorf, Germany). They were told to imagine that a close friend of them used to work in this department but was fired recently and got a very negative reference. The participants should search for this letter of reference and replace it with a better one. Moreover, they were told to look for a key to additionally find some valuables in the room that should be stolen. Participants had to read the instruction in detail before they were asked to describe the mock-crime in their own words to the assistant. In case they forgot some steps, they were corrected to ensure that the task was completely clear to them. During the course of the mock-crime, participants were confronted with 5 central and 5 peripheral details (printed in italics below) that were used as relevant items in the subsequent CIT examination. Importantly, none of the relevant details was explicitly mentioned in the instruction. Thus, all items were solely perceived and encoded during commitment of the mock-crime.

Participants had to search the storeroom by themselves and were instructed to mark the entry into this room in the heart rate recordings by pressing the marker button on the heart rate monitor watch. In case that other employees of the department spoke to them, they were told to react in an inconspicuous way and not to mention their involvement in the experiment. In the room, guilty participants first had to find a dolly where the letter of reference would be placed. Next to the dolly, a poster depicting a famous location in *Berlin* was placed. Participants had to move a *water box* to reach the dolly and were instructed to open all drawers. In the dolly, they found the letter of reference in a *folder* and in a different drawer a key with a *car* as key fob. In the other drawers, participants should additionally see some *paper-clips* and a CD, titled *Nautic Studie*. Subjects had to use the key to open a *cash box* placed at a shelf next to the drawer and to steal a 50 EUR *bank note*. Finally, subjects were instructed to put everything back in the original position, except for the money. While the subjects left the room they should see a *lamp* in one of the room corners and a *lab coat* hanging on the door. After completing the mock-crime they had to press the marker on the heart rate monitor watch after leaving the storeroom and again when returning to the examination room. At that time, the assistant asked participants to retrospectively rate their subjective nervousness during the mock-crime on the same Likert scale as before and to name any problems with respect to marker placement in the heart rate recordings. The answers were documented in a protocol and the subjects had to return the stolen money to the assistant.

Arousal was induced in one half of guilty participants by a confederate of the examiner who unexpectedly entered the storeroom shortly after the examinee arrived. The confederate was unknown to the examinees and he asked the participants if they were looking for something. Most participants replied with a short answer like “Yes” or just nodded and the confederate left the scenario quickly, so that participants could carry on. Such a simple question was used to allow the participants to respond very briefly without interrupting the whole scenario. It was important not to start a longer conversation to make sure that the situation was comparable for every subject. Participants in the condition without arousal induction were not interrupted during the mock-crime.

Innocent participants were instructed to imagine that they want to meet a person at the Department of Systems Neuroscience. Participants had to go to the storeroom, mark their arrival by pressing the respective button on the heart rate monitor watch and knock on the door. After getting no response, they should sit down on a chair placed close by in the corridor, wait for 3 minutes before setting another marker and returning to the examination room. Participants were asked to retrospectively rate their subjective nervousness during the task on the Likert scale.

Stage 2. After the mock-crime, a second experimenter, who was unknown to the examinees, conducted a CIT. For guilty examinees, this test was either carried out immediately after participants returned to the laboratory or 13-15 days later. The control group of innocents was tested immediately. Prior to the CIT, participants were told that the experiment was designed to check whether they could cope with the polygraph test and convince the examiner that they were innocent. A financial reward was offered for a successful performance in the CIT to increase motivation. Innocent participants could gain an extra reward of 5 EUR and guilty participants of 30 EUR for being classified as innocent.

Following these instructions, the examiner attached the polygraph devices. After adjustment of the eye tracking camera, a 9-point calibration procedure was completed before starting the CIT and repeated after half of the questions were presented. All participants were interrogated using a CIT with ten multiple-choice questions. An active wording of CIT questions was used (cf., Bradley, MacLaren, & Carle, 1996). Five of these questions focused on objects the participants had to actively handle during the mock-crime procedure and were therefore called central items (i.e., water box, folder, car key fob, cash box, bank note). The other five questions targeted peripheral items that were also placed in the storeroom but irrelevant for the execution of the mock-crime (i.e., Berlin picture, paper-clips, Nautic Studie CD, lamp, lab coat). Each CIT question (e.g., “What did you steal from the room?”) consisted of one buffer item (e.g., a projector) following question presentation, four irrelevant items (e.g., a mobile phone, an external hard disk, a camera, a laptop) and one crime-related relevant item (e.g., a bank note). All questions and items are reported in Table 1.

Table 1

Questions and items used in the Concealed Information Test.

Question	Item 1 (buffer)	Item 2	Item 3	Item 4	Item 5	Item 6
What did you move to reach the dolly?	Toiletpaper	Tool box	Box of paper	Printer cartridge	Water bottles*	Waste paper basket
Which city did you see on the poster hanging in the room?	Pisa	London	New York	Berlin*	Paris	San Francisco
Where were the documents you had to destroy?	Punched pocket	Envelope	Report folder	Folder*	Binder	Cardboard file folder
What was written on the CD that you saw in the dolly?	Schmerz Studie	Alters Studie	Sisal Studie	Effekt Studie	Nautic Studie*	Memory Studie
What kind of key fob was attached to the key you found?	Knife	Car*	Cord	Rose	Lock	Ribbon
What was placed next to the key you found in the dolly?	Tape	Pencil	Notes	Ball-pen	Marker	Paper-clips*
Where was the valuable thing placed that you stole?	Safe	Suitcase	Wall cupboard	Wooden box	Bag	Cash box*
What did you steal from the room?	Projector	Mobile phone	Money*	External hard disk	Camera	Laptop
What did you see hanging on the door?	Work overall	Lab coat*	Plastic bag	Multiple socket	Umbrella	Pouch
What else did you see in the room?	Chairs	Screen	Lamp*	Pail	Folding bed	Ladder

Note: Questions and items are loosely translated in the present table. It was ensured that the original details were easy to discriminate by the subjects. * = the relevant items.

The order of relevant and irrelevant items within each question was randomly determined but remained constant across subjects. The CIT questions were asked in a fixed order representing the sequence of actions during the mock-crime. Central and peripheral questions alternated with the exception of the last four questions, where two peripheral questions followed two central ones. Each trial started with a central fixation cross that appeared on the screen for 2 s. Then, a prerecorded audio sample naming the item was presented while a photograph of the item was shown for 5 s. Afterwards, a blank screen (uniform gray background) followed for approximately 13 s. Participants were instructed to look at the fixation cross at trial start and to respond by verbally answering “No” to every item (see Figure 2). The presented pictures had a size of approximately 12.5° by 12.5° of visual angle, depending on the exact distance between participant and screen. The CIT was separated into two blocks, with a short break after five questions.

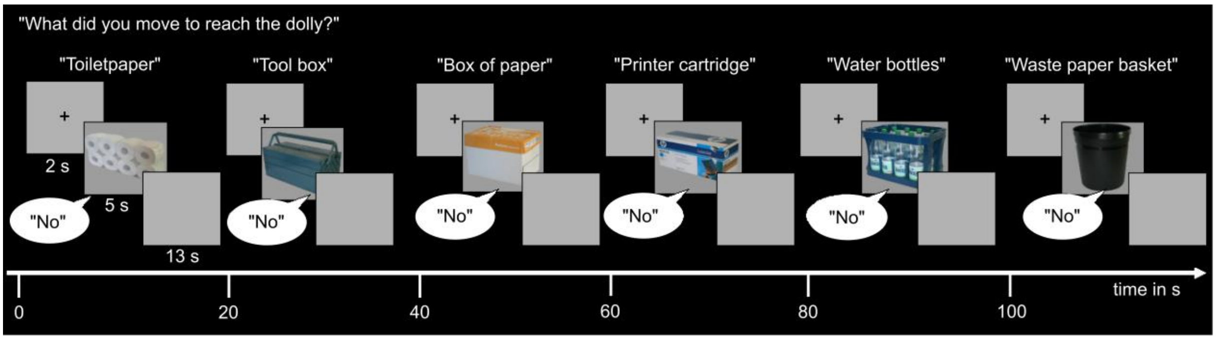


Figure 2. Event sequence during the CIT (Study I).

After the CIT was finished, the polygraph devices were detached and participants were asked to reveal whether they committed the mock-crime. Subsequently, an explicit recognition test was accomplished. For this test, all CIT questions were presented sequentially along with all corresponding items. Participants were asked to select the correct detail. If they forgot the information, they were instructed to guess. Finally, all participants were debriefed and were gently requested not to talk to other possible participants about the procedure and content of the study.

2.2.5 Response Scoring and Analysis

Data of participants within the innocent condition were only used to calculate validity coefficients. However, results on heart rate and subjective nervousness ratings during the experiment are reported to show that the experimental task was much less arousing than the mock-crime. The innocent group was not included in the analyses on experimental manipulations, which only focus on differences between the guilty groups. The same is done for the CIT data because since innocents did not encode the relevant details of the crime, no effect on CIT results can be expected.

Heart rate and subjective nervousness during the mock-crime. Data of the portable heart rate monitor were checked for marker placement and corrected if necessary, based on study protocols. Due to equipment failure, data of one guilty subject were lost for the heart rate analysis. The statistical programming language R (www.r-project.org) was used to implement the following calculations. After correction of artifacts, specific phases of the mock-crime were isolated and mean heart rate within these phases was calculated. The baseline period had a fixed length of 120 s. During the first experimental phase, participants searched for the storeroom (mean duration = 72.5 s, *SD* = 35.2 s). The next phase corresponded to the time subjects spent in the storeroom (mean duration = 231.1 s, *SD* = 90.3 s) and during the last phase, they were heading back to the examination room (mean duration = 60.8 s, *SD* = 25.7 s). On a single subject level, the baseline mean value was subtracted from the mean value of all three experimental phases, respectively, and these baseline corrected data were used for further analyses. Success of arousal manipulation was tested by a 2 × 2 × 3 mixed model analyses of variance (ANOVA) using the between subject factors arousal manipulation (arousal induction vs. no arousal induction) and delay between mock-crime and CIT (immediately vs. two weeks delayed) and the within-subjects factor phase of mock-crime (baseline corrected heart rate of phases one to three). To circumvent the potentially problematic sphericity assumption of univariate analyses involving repeated measures, a multivariate approach was used here.

A similar procedure was realized with the subjective nervousness ratings that were given on the Likert scale ranging from 1 (*not nervous at all*) to 6 (*very nervous*). Baseline corrected scores for the ratings delivered before (phase one) and during the mock-crime (phase two) were calculated and tested by a 2 × 2 × 2 mixed model ANOVA with the between subject factors arousal manipulation and delay between mock-crime and CIT and the within-subjects factor phase of mock-crime (baseline corrected ratings of phases one and two). Comparable analyses were conducted for the group of innocent participants to determine whether subjective nervousness or heart rate varied across the different phases of their task.

Recognition task. The number of correctly recognized relevant details was calculated for every participant and the effect of the experimental manipulations was tested by means of a 2 × 2 × 2 mixed model ANOVA using the between subject factors arousal manipulation and delay between mock-crime and CIT and the within-subjects factor type of critical detail (central vs. peripheral).

Physiological responses in the CIT. Physiological responses in the CIT were scored similarly to previous studies (e.g., Gamer et al., 2006, 2010). The statistical programming language R was used to implement all calculations except for heart rate responses that additionally required visual inspection.

For electrodermal responses, the amplitude of the largest skin conductance increase that occurred between 0.5 and 10.5 s after item onset was calculated automatically. Zero-responses to single items as well as participants with a large number of such responses (i.e., electrodermal non-responders) were included in the analyses to be consistent with previous studies (e.g., Gamer et al., 2006, 2010). Respiration line length (RLL) was scored during the interval 0-10 s following item onset. For the analysis of heart rate responses, R-waves were detected from the ECG data. The automatically detected R-waves were visually inspected and

R-R-intervals were converted to HR (in beats per minute). Afterwards, a second-by-second sampling was applied (Velden & Wölk, 1987) resulting in one HR value for each of 15 poststimulus seconds. The HR in the last second prior to item onset represented the prestimulus baseline. Poststimulus difference scores (ΔHR) were derived by subtracting the prestimulus baseline value from the HR-score of each poststimulus second. Finally the average of all ΔHR values was calculated (cf., Gamer, Verschuere, et al., 2008; Verschuere et al., 2004). FPWL was calculated analogously as RLL during the interval 0-15 s following item onset (Elaad & Ben-Shakhar, 2006).

To facilitate a comparison between measures and examinees, standard difference scores were calculated from the physiological responses according to a procedure described by Gamer et al. (2010). Within each participant, the difference between the critical item and the mean of all irrelevant items within one question was calculated and this difference was divided by the standard deviation of all irrelevant items of the test. Thus, the response differences for each question were calculated and scaled with the individual response distribution of the irrelevant items within each examinee. The mean of these measures was computed for central and peripheral details separately. Depending on the measure, a positive (electrodermal response) or negative value (respiration, finger pulse and HR) indicates recognition of relevant crime details. Values around zero suggest that the participant could not differentiate between relevant and irrelevant details which is the expected response pattern for innocents. A combination of the physiological measures (excluding finger pulse) was calculated based on a previously proposed and validated logistic regression equation (Gamer, Verschuere et al., 2008). The regression coefficients used for calculation of the combined classification score were $\beta = -3.917$ (constant), $\beta = 4.241$ (electrodermal response), $\beta = -6.310$ (respiration) and $\beta = -1.975$ (heart rate). This classification function was adopted separately for central and peripheral details as well as for the whole test. The effect of the experimental manipulations on the standard difference scores as well as on the combined classification score was tested by means of separate 2 × 2 × 2 mixed model ANOVAs using the between subject factors arousal manipulation and delay between mock-crime and CIT and the within-subjects factor type of critical detail. For all analyses, the a priori significance level was set to $\alpha = .05$ but marginally significant results are reported when $p < .10$. Cohen's f is reported as an effect size estimate (Cohen, 1988, p. 273ff).

Eye-movement responses in the CIT. After blink detection, eye movement data were parsed into saccades and fixations using Eyelink's standard parser configuration, which classifies an eye movement as a saccade when it exceeds 30°/sec velocity or 8,000°/sec^2 acceleration. Time intervals between saccades were defined as fixation. Subsequently, x and y coordinates of fixations were drift corrected with reference to the central fixation cross at the start of each trial. All pictures had the same size, therefore a central screen area of 600 × 600 pixel was defined as region of interest (ROI). Fixations were attributed to an item when they were within the region's pixel coordinates. Three measures were derived from the eye movement recordings: the number of fixations on the ROI, the average duration of these fixations in ms and the number of blinks. All measures were calculated for three different time intervals of 5 s each starting with stimulus onset (0 s) and ending 15 s after stimulus onset. Therefore, it was able to analyze eye tracking data during stimulus presentation (first interval) as well as after stimulus offset (second and third interval). Similar to the autonomic data, each subject's raw values were converted to standard difference scores to facilitate a comparison between measures and examinees. The mean of these measures was computed separately for central and peripheral details as well as for all three time intervals. Since less blinks and fewer fixations on relevant details were expected when examinees were able to identify the critical item, standard difference scores should be negative for guilty examinees. Furthermore, increased fixation durations should result in positive standard difference scores. In contrast, measures around zero were predicted for innocents who should show a non-systematic response pattern to relevant and neutral CIT items.

One guilty and one innocent subject were excluded from the blink data due to missing blinks. For all guilty subjects a 2 × 2 × 2 × 3 analysis of variance (ANOVA) was calculated for each ocular measure with the between subject factors arousal manipulation (arousal induced vs. not induced) and time between mock-crime and CIT (immediately vs. 2 weeks later) and the within subject factors type of critical detail (central vs. peripheral) and time interval after stimulus onset (0-5 s, 5-10 s, 10-15 s). Cohen's *f is* reported as effect size estimates (Cohen, 1988, p. 273ff). For statistically significant post-hoc analysis Cohens *d* is reported (Cohen, 1988, p. 19ff).

To investigate the interrelation of ocular measures, bivariate pearson's correlations were calculated. Since electrodermal, respiratory and cardiovascular measures were additionally acquired during the CIT, the relationship between eye tracking measures and autonomic responses was also examined. For this purpose, each ocular measure was correlated with a combination score calculated for the autonomic responses across all guilty examinees. Correlations were determined separately for central and peripheral details.

Finally, it was investigated whether ocular measures provide incremental validity to autonomic responses in differentiating guilty and innocent examinees. Because of the small sample size of individual groups (n = 17) and the much smaller number of the innocent examinees (n = 17) as compared to all guilty subjects (n = 68), a bootstrapping approach was used for this purpose. Thus, a sample of n = 17 innocents and n = 17 guilty examinees was drawn (with replacement) from the respective group of innocents and one group including all guilty subjects. These data were used to calculate three logistic regression models with the combined physiological score (cf., Gamer, Verschuere, et al., 2008), the combined ocular score, or both measures as predictors. For the combined ocular score, standard difference scores for the number of fixations in the time interval from 0 to 5 s and the number of blinks in the interval from 5 to 10 s were averaged. These two ocular measures were chosen as they showed the highest validity coefficients and a comparably small intercorrelation (see below).

The logistic regression models were only calculated for central details since ocular measures did not differ significantly between relevant and neutral CIT items for peripheral details (see below). The coefficients of the logistic regression models and Nagelkerke's R^2 as a measure of the overall effect size were saved and the whole procedure was repeated 10,000 times for different random samples of guilty and innocent examinees. Samples in which the estimation of the logistic function did not converge within 25 iterations were excluded. From the stored data, mean regression coefficients and Nagelkerke's R^2 were calculated for each model. Moreover, p-values for each coefficient were derived by calculating the proportion of bootstrap samples that fell below (for positive coefficients) or above 0 (for negative coefficients). These values were multiplied by 2 to implement a two-sided test.

2.3 Results

2.3.1 Manipulation checks

The 2 × 2 × 2 mixed model ANOVA on baseline corrected subjective nervousness ratings revealed a significant interaction of arousal induction and phase of mock-crime, $F(1, 64) = 6.97$, $p = .010$, $f = 0.16$, as well as a main effect of phase of mock-crime, $F(1, 64) = 31.34$, $p < .001$, $f = 0.35$. The 2 × 2 × 3 mixed model ANOVA on baseline corrected heart rates also revealed a significant interaction of arousal induction and phase of mock-crime, $F(2, 62) = 3.44$, $p = .038$, $f = 0.10$, as well as a main effect of phase of mock-crime, $F(2, 62) = 8.65$, $p < .001$, $f = 0.11$. As shown in Figure 3, all guilty groups showed an increase in their subjective nervousness rating when committing the mock-crime in comparison to their ratings before the mock-crime. This effect was stronger in the arousal induction group. For heart rate changes, an increase relative to baseline was detected in all three phases of the experiment. This increase was most prominent in the phase after the commitment of the mock-crime, especially in the arousal induction group.

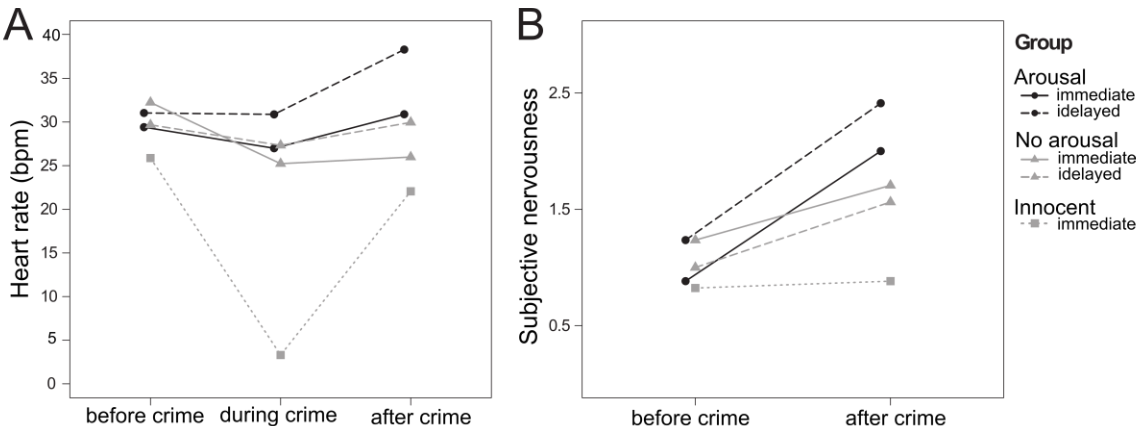


Figure 3. A) Mean heart rate and B) subjective nervousness ratings (Study I). Depicted are different phases of the mock-crime experiment as a function of arousal induction and time of CIT examination. The innocent group is additionally depicted for comparison. All values show the mean value minus the relevant baseline value. Reprinted with permission from Peth, J., Vossel, G., & Gamer, M. (2012). Emotional arousal modulates the encoding of crime-related details and corresponding physiological responses in the Concealed Information Test. *Psychophysiology*, *49*(3), 381–390.

For the innocent group, no arousal increase was found during the relevant phase of the experiment. The within subject ANOVA on baseline corrected subjective nervousness scores revealed no effect for the factor phase of experiment, $F(1, 16) < 1$, $f = 0.03$. The ANOVA on baseline corrected heart rate scores revealed a significant effect for phase of experiment, $F(2, 15) = 128.14$, $p < .001$, $f = 1.13$. As shown in Figure 3A, this effect resulted from a relative heart rate decline during the second phase of the experiment when innocent participants sat down and waited.

2.3.2 Memory test

The 2 × 2 × 2 mixed model ANOVA on the number of correctly recognized items revealed a significant interaction of arousal induction and type of critical detail, $F(1, 64) = 5.92$, $p = .018$, $f = 0.12$, as well as significant main effects of arousal induction, $F(1, 64) = 7.73$, $p = .007$, $f = 0.13$, time of test, $F(1, 64) = 12.78$, $p = .001$, $f = 0.17$, and type of critical item, $F(1, 64) = 259.22$, $p < .001$, $f = 1.33$. As can be seen in Figure 4, participants in the arousal induction group and in the no arousal induction group did not differ in their recognition of central details. Both groups showed a reduced recognition of peripheral items, which was more pronounced for participants in the arousal induction group. A general decline in the amount of correctly recognized items from the immediate to the delayed test was found.

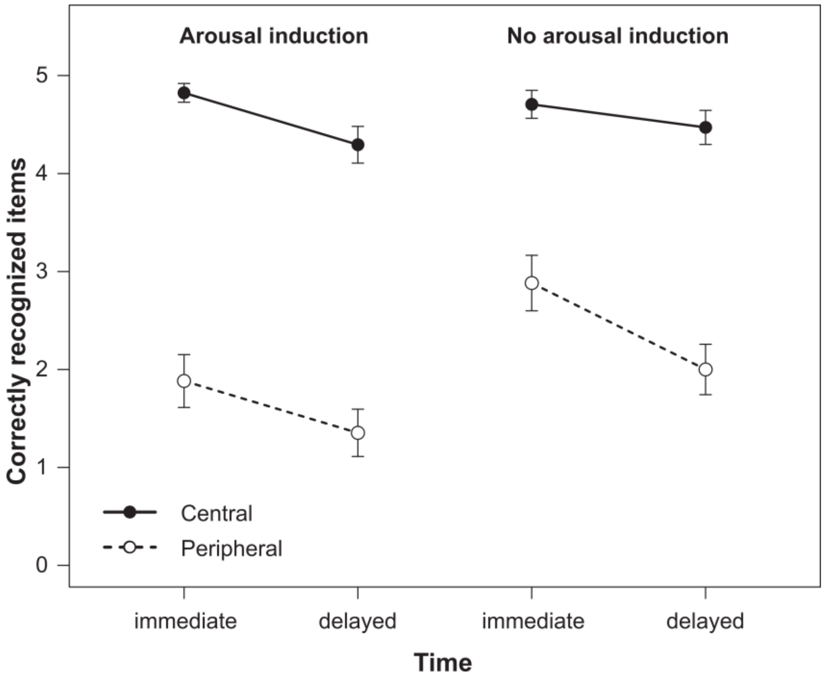


Figure 4. Number of remembered items in the recognition test (Study I) as a function of arousal induction, time of CIT examination, and type of relevant detail, Overall, there were five central and five peripheral details that served as relevant items in the CIT. Error bars indicate standard errors of the mean. Reprinted with permission from Peth et al., 2012.

2.3.3 Physiological measures

The separate 2 × 2 × 2 mixed model ANOVAs on the physiological responses in the CIT revealed a statistically significant main effect of type of critical detail on the electrodermal response, $F(1, 64) = 63.85$, $p < .001$, $f = 0.59$, on the respiratory response, $F(1, 64) = 22.06$, $p < .001$, $f = 0.36$, on the finger pulse response, $F(1, 64) = 28.17$, $p < .001$, $f = 0.46$ and on the heart rate response, $F(1, 64) = 14.55$, $p < .001$, $f = 0.32$. As depicted in Figure 5, all measures showed larger response differences between relevant and irrelevant items for central as compared to peripheral details.

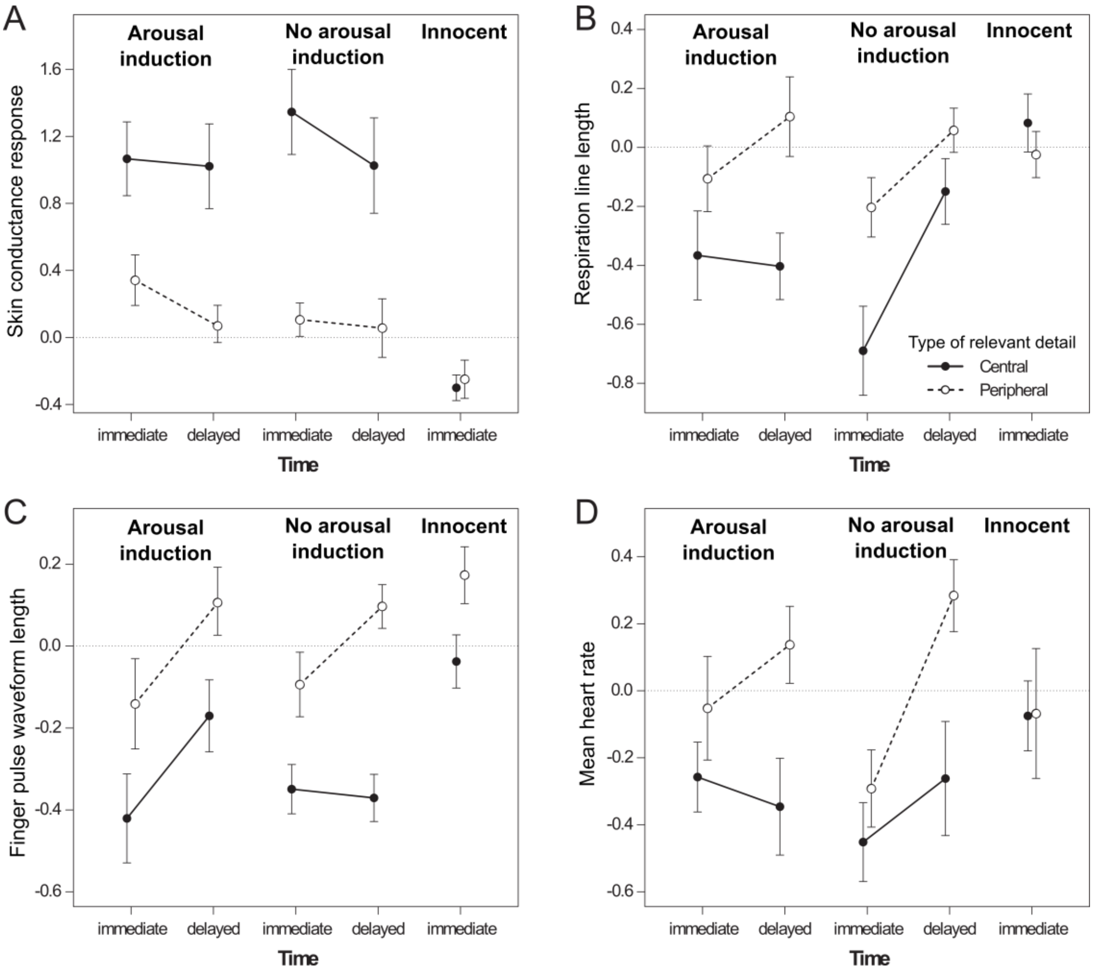


Figure 5. Standard difference scores of the responses to relevant and irrelevant items (Study I) for skin conductance responses (A), respiration line length (B), finger pulse waveform length (C) and mean heart rate (D) as a function of group, time of CIT examination and type of relevant detail. Error bars indicate standard errors of the mean. Recognition of relevant details is reflected by positive values for skin conductance responses and negative values for respiratory, finger pulse and heart rate measures. Error bars indicate standard errors of the mean. Reprinted with permission from Peth et al., 2012.

A significant main effect of time of test was found on the respiratory response, $F(1, 64) = 6.90$, $p = .011$, $f = 0.23$, on the finger pulse response, $F(1, 64) = 8.45$, $p = .005$, $f = 0.22$ and on the heart rate response, $F(1, 64) = 5.38$, $p = .024$, $f = 0.19$. Thus, with exception of the electrodermal response, all measures tended to show a decreased differentiation between relevant and irrelevant items at the delayed measurement occasion (Figure 5). However, this effect was more pronounced in the condition without arousal induction especially with respect to respiratory and heart rate responses (see Figure 5B and 5D). Consequently, we additionally observed trends for an interaction of arousal induction and time of test for these measures: $F(1, 64) = 2.87$, $p = .095$, $f = 0.15$ (for respiratory responses), $F(1, 64) = 3.17$, $p = .080$, $f = 0.15$ (for heart rate responses).

Furthermore, a trend for an interaction of time of test and type of critical item on the heart rate responses, $F(1, 64) = 3.29$, $p = .074$, $f = 0.15$, showed that central details tended to elicit temporally more stable response differences than peripheral details. Finally, a trend for a three-way interaction of arousal induction, time of test and type of critical detail on the respiratory responses, $F(1, 64) = 2.86$, $p = .096$, $f = 0.12$ was found, pointing to a more stable reaction over time for central details within the arousal induction group (Figure 5B).

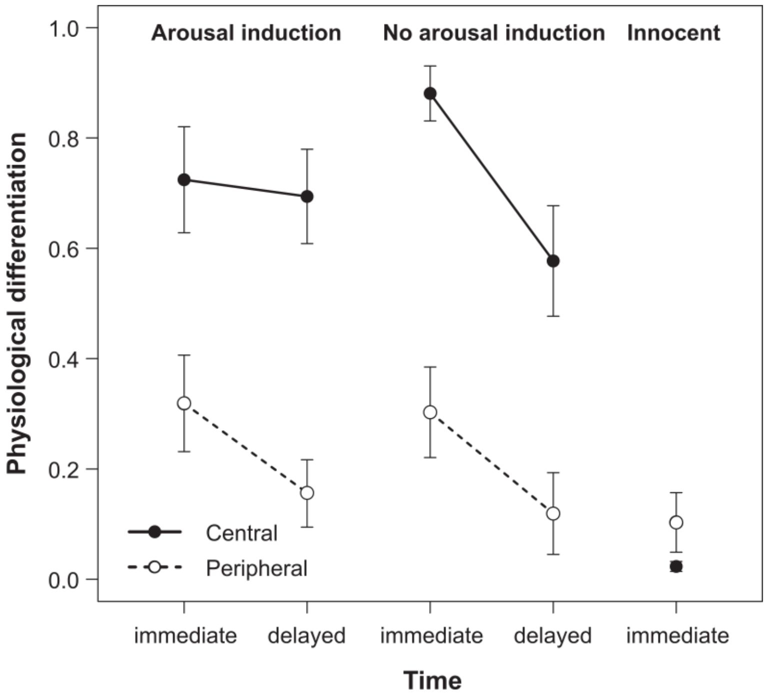


Figure 6. Strength of physiological differentiation between relevant and irrelevant CIT items (Study I) as a function of arousal induction, time of CIT examination, and type of relevant detail. This score reflects the combination of the standardized response differences of electrodermal, respiratory and heart rate data by means of a previously described logistic regression function (Gamer, Verschuere, et al., 2008). A value of 1 indicates perfect physiological differentiation of relevant and irrelevant items whereas a value of 0 reflects unsystematic responding to both item types. Error bars indicate standard errors of the mean. Reprinted with permission from Peth et al., 2012.

With respect to the combined classification score of electrodermal responses, respiration and heart rate, a 2 × 2 × 2 ANOVA only revealed statistically significant main effects of time of test, $F(1, 64) = 7.30$, $p = .009$, $f = 0.21$, and type of critical detail, $F(1, 64) = 93.80$, $p < .001$, $f = 0.73$. In general, a stronger differentiation of relevant and irrelevant CIT items for central as compared to peripheral details was found as well as an advantage of immediate over delayed testing (see Figure 6).

Validity of the CIT with autonomic measures

The above-mentioned analyses mainly focused on factors influencing the response pattern of guilty examinees in the CIT. Thus, these analyses are not directly suitable to examine the validity of the test in differentiating guilty and innocent subjects. For this aim, an additional group of innocent participants was examined. To determine the degree of differentiation between guilty and innocent participants, separate receiver operating characteristics (ROC) curves were calculated, contrasting the combined physiological classification score of all four guilty groups (defined by the factors arousal induction and time of test) with the group of innocents. The area under the ROC curve can be interpreted as a validity coefficient (Ben-Shakhar & Elaad, 2003). An area of 1 indicates perfect differentiation of the respective groups whereas a value of 0.5 reflects chance classification. In addition to this descriptive value, 95% confidence intervals were calculated to examine whether the area statistics exceed the value 0.5 and thus indicate a significant differentiation from innocent subjects (Bamber, 1975).

As shown in Figure 7, a CIT based on the physiological responses to all questions allowed for a clear differentiation from innocent participants for both guilty groups (arousal induction and no arousal induction) and over both measurement occasions (immediate and delayed test). The no arousal induction group showed a slightly better differentiation for the immediate CIT, but the area statistic declined significantly over time, $z = 2.69$, $p = .007$, which was not observed in the arousal induction group, $z = 1.19$, $p = .234$ (see Hanley & McNeil, 1983; McNeil & Hanley, 1983; for the calculations). Separate calculation for central and peripheral details revealed different results for the two types of details. Peripheral questions showed a decline over time within both groups and the 95% confidence intervals included chance level for the delayed CIT examination. In contrast, the scores for central questions remained above chance level for both groups. Interestingly, the area statistic remained temporally stable in the arousal induction group, $z = 0.85$, $p = .395$. By contrast, in the no arousal induction group the area statistic reached a score of 1 at the immediate CIT application, but it declined significantly over time, $z = 2.01$, $p = .045$.

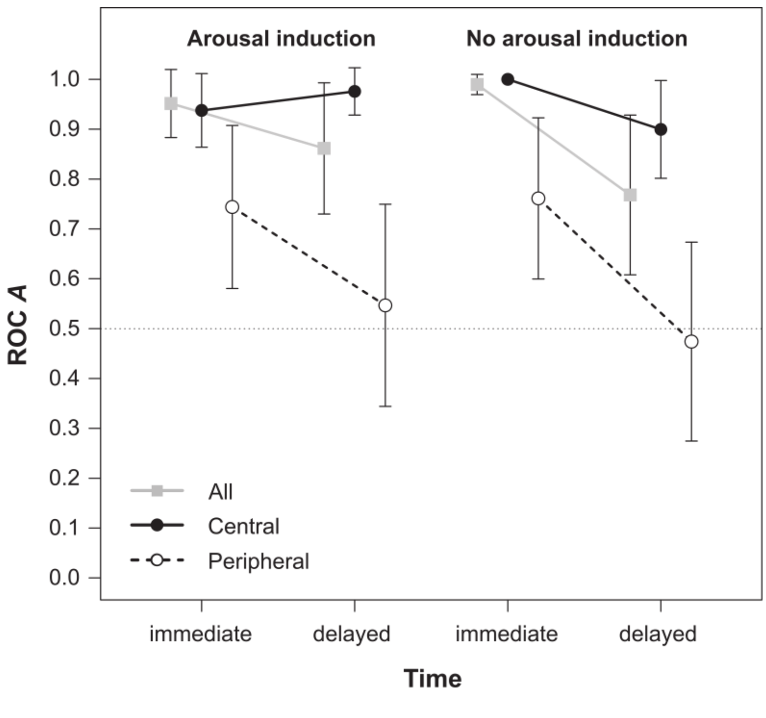


Figure 7. Areas under the ROC curves (Study I) as a function of arousal induction, time of CIT examination, and type of relevant detail. The ROC curves rely on the combined classification score using the group of innocents as the basis for comparison. Area values close to 1 reflect a high differentiation of the respective group from the innocents, whereas values close to 0.5 indicate chance classification. Error bars indicate corresponding 95% confidence intervals. Reprinted with permission from Peth et al., 2012.

2.3.4 Eye-movements

Categorial analyses of fixations and blinks

The ANOVA on the number of fixations revealed significant main effects of the intercept, $F(1,64) = 5.93$, $p = .018$, $f = 0.26$, time of test, $F(1,64) = 4.58$, $p = .036$, $f = 0.22$, type of critical detail, $F(1,64) = 12.23$, $p < .001$, $f = 0.32$, and time interval after stimulus onset, $F(2,63) = 4.96$, $p = .010$, $f = 0.23$. Furthermore, a significant 4-way interaction was found between all factors, $F(2,63) = 4.31$, $p = .018$, $f = 0.23$. All remaining effects failed to reach statistical significance (all $p > .076$, $f < 0.19$).

To decompose the 4-way interaction, separate 3-way ANOVAs were calculated for each time interval. For the first time interval (0-5 s) the ANOVA revealed significant main effects of the intercept, $F(1,64) = 7.60$, $p = .008$, $f = 0.48$, the time of test, $F(1,64) = 4.84$, $p = .031$, $f = 0.37$, and the type of critical detail, $F(1,64) = 17.08$, $p < .001$, $f = 0.59$. Furthermore, a significant 3-way interaction was found between time of test, arousal induction and type of critical detail, $F(1,64) = 5.95$, $p = .017$, $f = 0.31$. The remaining effects were not statistically significant (all $p >.106$, $f < 0.27$). As shown in Figure 8A, the difference between central and peripheral details during stimulus presentation increased from the immediate to the delayed measurement occasion within the arousal induction group. By contrast, the group without arousal induction showed no such change. The ANOVA for the second time interval (5-10 s) revealed significant main effects for the intercept, $F(1,64) = 6.99$, $p = .010$, $f = 0.40$, and type of critical detail, $F(1,64) = 4.05$, $p = .048$, $f = 0.29$. All remaining effects did not reach statistical significance (all $p > .127$, $f < 0.22$). Thus, in the time interval directly following stimulus offset, only a stronger effect for central compared to peripheral details was found. Finally, the ANOVA for the third time interval (10-15 s) revealed no statistically significant effects (all $p > .203$, $f < 0.16$).

With respect to the other results obtained in the main analysis, the significant intercept indicates that the number of fixations on relevant details differed from that on neutral items across all experimental factors. Moreover, as shown in Figure 8A, the number of fixations on central relevant details was smaller compared to peripheral ones and overall, effects decreased across time intervals and between the immediate and the delayed test.

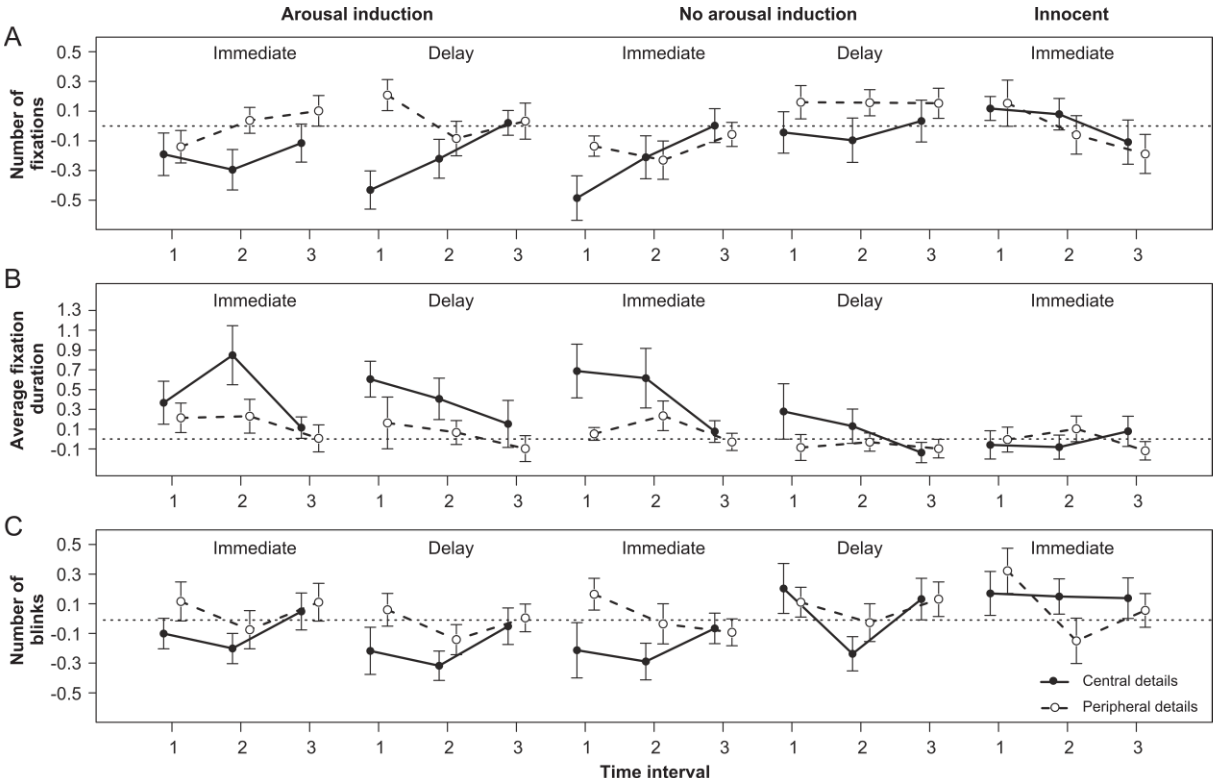


Figure 8. Differentiation between relevant and irrelevant CIT items with respect to the number of fixations (panel A), the average fixation duration (panel B) and the number of blinks (panel C) in Study I. Depicted as a function of arousal induction, time of CIT examination, and type of relevant detail. During time interval 1 (0 – 5 s after stimulus onset), the stimuli were presented as photographs on the screen. Time intervals 2 (5 – 10 s) and 3 (10 – 15 s) correspond to the time after stimulus offset where a blank screen was shown. Negative values indicate less fixations, shorter fixations and less blinks on relevant details whereas a value of 0 reflects unsystematic responding to relevant and irrelevant CIT items. Error bars indicate standard errors of the mean. Reprinted with permission from Peth, J., Kim, J. S.-C., & Gamer, M. (2013). Fixations and eye-blinks allow for detecting concealed crime related memories. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *88*(1), 96–103.

The ANOVA on the average fixation duration revealed significant main effects of the intercept, $F(1,64) = 20.84$, $p < .001$, $f = 0.33$, type of critical detail, $F(1,64) = 12.37$, $p = .001$, $f = 0.23$, and time interval after stimulus onset, $F(2,63) = 7.74$, $p = .001$, $f = 0.26$. All remaining effects were not statistically significant (all $p > .053$, $f < 0.13$). The significant intercept indicates that the average fixation duration differed between relevant and neutral CIT items irrespective of the experimental manipulations. Furthermore, as shown in Figure 8B, the average fixation duration on central relevant details was larger compared to peripheral ones. Finally, standard difference scores declined across time and approximated zero in the third interval (10-15 s).

The ANOVA on the number of blinks revealed significant main effects of type of critical detail, $F(1,63) = 5.07$, $p = .028$, $f = 0.13$, and time interval after stimulus onset, $F(2,62) = 8.74$, $p < .001$, $f = 0.17$, as well as a significant interaction between arousal manipulation and time of test, $F(1,63) = 5.09$, $p = .028$, $f = 0.11$. All remaining effects did not reach statistical significance (all $p > .117$, $f < 0.08$). To decompose the interaction effect, single t-tests were calculated between the arousal conditions for the immediate and the delayed CIT. The results revealed no differences between the groups with and without arousal induction at the immediate test, $t(31) < 1$. However, a significant difference was observed for the delayed testing occasion, $t(32) = 2.18$, $p = .036$, $d = 0.75$, where larger response differences between relevant and neutral CIT items were discovered for the arousal induction group. As shown in Figure 8C, less blinks occurred for central relevant details compared to peripheral ones and overall, effects were most prominent in the time interval immediately after stimulus offset (5-10 s). This latter finding was further supported by a post-hoc ANOVA for the time interval from 5 to 10 s. Only this time interval revealed a significant main effect for the intercept, $F(1,63) = 17.40$, $p < .001$, $f = 0.37$. Additionally, the type of critical detail was significant, $F(1,63) = 4.97$, $p = .029$, $f = 0.21$, while all remaining effects did not reach statistical significance (all $p > .447$, $f < 0.06$).

Correlative analyses of ocular and autonomic measures

First, Pearson's correlations were calculated between standard difference scores of all eye tracking measures for each of the three time intervals (0-5 s, 5-10 s, 10-15 s) across all guilty examinees. As shown in Table 2, the number of fixations and the average fixation durations were strongly correlated during all time intervals. This relationship was found for central and peripheral details and indicates that fewer fixations on relevant items were accompanied by an enhanced fixation duration. By contrast, only moderate negative correlations were found between the average fixation durations and the number of eye-blinks. This effect was restricted to central details and largest in the time interval immediately following stimulus offset. For the number of fixations, moderate positive correlations with the number of eye-blinks were observed in the 5 – 10 s time interval after presentation of central and peripheral details.

Table 2.

Correlations between ocular measures (number of fixations, average fixation duration and number of blinks) across all guilty examinees.

	Time interval		
Measures	0 – 5 s	5 – 10 s	10 -15 s
Number of fixations x Average fixation duration	(*n* = 68)	(*n* = 68)	(*n* = 68)
Central details	-0.812***	-0.591***	-0.297*
Peripheral details	-0.562***	-0.513***	-0.498***
Number of fixations x Number of blinks	(*n* = 67)	(*n* = 68)	(*n* = 68)
Central details	0.219	0.325**	0.114
Peripheral details	0.016	0.310*	-0.032
Average fixation duration x Number of blinks	(*n* = 67)	(*n* = 68)	(*n* = 68)
Central details	-0.289*	-0.383**	-0.257*
Peripheral details	-0.064	-0.222	0.075

Note. * $p < .05$, ** $p < .01$, *** $p<.001$.

In a second step, Pearson's correlations were calculated between ocular measures and the autonomic physiological responses that were additionally recorded during the CIT. Correlations were computed between standard difference scores of the number of fixations, the average fixation duration and the number of blinks with a combination score of electrodermal, respiratory and heart rate responses for each of the three time intervals (0-5 s, 5-10 s, 10-15 s).

As shown in Table 3, significant negative correlations were obtained between the number of fixations and the combined physiological score during the first two time intervals. For the average fixation duration, a positive correlation was observed during the second time interval. These effects were restricted to central crime details. Thus, larger autonomic responses to such crime-related items were accompanied by a reduced number of fixations and prolonged fixation durations on these details. For the number of eye-blinks no significant correlation with the combined physiological score was found.

Table 3.

Correlations between ocular measures (number of fixations, average fixation duration and number of blinks) and a combined physiological score (electrodermal, respiratory and heart rate measures) across all guilty examinees.

		Time interval	
Measure	0 – 5 s	5 – 10 s	10 -15 s
Number of fixations	(n = 68)	(n = 68)	(n = 68)
Central details	-0.337**	-0.307*	-0.165
Peripheral details	-0.060	-0.112	-0.075
Average fixation duration	(n = 68)	(n = 68)	(n = 68)
Central details	0.195	0.241*	0.074
Peripheral details	0.059	0.059	-0.027
Number of blinks	(n = 67)	(n = 68)	(n = 68)
Central details	-0.128	-0.032	0.020
Peripheral details	0.114	0.001	-0.143

Note: * p < .05, ** p < .01.

Validity of the CIT with ocular measures

In addition to the four guilty groups, an innocent group was tested to allow for estimating the validity of the CIT in differentiating guilty and innocent subjects. To this aim, separate receiver operating characteristics (ROC) curves were calculated for the time intervals from 0-5 s and 5-10 s, contrasting the ocular measures in each guilty group with the group of innocents. The last time interval from 10-15 s was excluded from this analysis since Figure 8 shows that standard difference scores approximate zero, which indicates that the eye tracking data did not differ between relevant and neutral CIT items in this time interval. Additionally, ROC curves were calculated across the factors arousal manipulation and time of test as well as for a group of all guilty subjects.

For the fixation data, a valid differentiation between guilty and innocent subjects was only possible when relying on central details. As shown in Table 4, only two (number of fixations) or three (average fixation duration) out of eight comparisons exceeded chance level. Validity coefficients seemed to be largest during stimulus presentation (0-5 s after stimulus onset) and were enhanced for the delayed test of the arousal induction group as well as the immediate CIT of the group without arousal induction. Additionally, a valid group differentiation was possible based on the average fixation duration in the time interval after stimulus offset (5-10 s after stimulus onset) in the immediate CIT of the group with arousal induction.

For groups pooled over arousal, areas under the ROC curve of both time intervals reached statistical significance during the immediate as well as the delayed CIT session. Only the average fixation duration after stimulus offset did not enable a valid classification for a delayed CIT.

Table 4.

Areas under the ROC curves (and 95 % confidence intervals) for the fixation data.

	Arousal induction		No arousal induction		Across arousal	
Time after stimulus:	0 – 5 s	5 – 10 s	0 – 5 s	5 – 10 s	0 – 5 s	5 – 10 s
Number of fixations:						
Immediate CIT						
Central details	0.67 (0.47,0.87)	0.68 (0.49,0.82)	**0.82 (0.66,0.97)**	0.66 (0.47,0.85)	**0.71 (0.57,0.87)**	**0.77 (0.65,0.90)**
Peripheral details	0.63 (0.43,0.82)	0.60 (0.39,0.80)	0.62 (0.40,0.83)	0.61 (0.41,0.80)	0.45 (0.28,0.63)	0.51 (0.33,0.68)
Delayed CIT						
Central details	**0.80 (0.64,0.95)**	0.63 (0.44,0.82)	0.59 (0.39,0.78)	0.57 (0.38,0.77)	**0.72 (0.57,0.87)**	**0.67 (0.52,0.82)**
Peripheral details	0.55 (0.35,0.75)	0.49 (0.29,0.69)	0.55 (0.35,0.75)	0.63 (0.44,0.82)	0.51 (0.34,0.67)	0.49 (0.31,0.67)
Across time of test					**Across arousal and time of test**	
Central details	**0.79 (0.66,0.92)**	**0.74 (0.60,0.88)**	0.64 (0.48,0.79)	**0.70 (0.55,0.84)**	**0.72 (0.58,0.85)**	**0.72 (0.60,0.84)**
Peripheral details	0.49 (0.32,0.66)	0.53 (0.34,0.71)	0.46 (0.29,0.64)	0.47 (0.29,0.64)	0.47 (0.32,0.63)	0.50 (0.33,0.66)
Average fixation duration:						
Immediate CIT						
Central details	0.63 (0.44,0.82)	**0.77 (0.61,0.93)**	**0.75 (0.58,0.93)**	0.66 (0.48,0.85)	**0.69 (0.54,0.84)**	**0.72 (0.58,0.86)**
Peripheral details	0.60 (0.41,0.80)	0.55 (0.35,0.75)	0.62 (0.43,0.82)	0.56 (0.36,0.75)	0.61 (0.44,0.79)	0.55 (0.38,0.73)
Delayed CIT						
Central details	**0.75 (0.59,0.92)**	0.64 (0.45,0.83)	0.57 (0.37,0.77)	0.58 (0.38,0.78)	**0.66 (0.51,0.82)**	0.61 (0.45,0.77)
Peripheral details	0.51 (0.31,0.71)	0.52 (0.32,0.71)	0.55 (0.35,0.75)	0.56 (0.37,0.76)	0.48 (0.31,0.65)	0.54 (0.36,0.72)
Across time of test					**Across arousal and time of test**	
Central details	**0.69 (0.54,0.84)**	**0.70 (0.56,0.85)**	**0.66 (0.51,0.82)**	0.62 (0.47,0.78)	**0.67 (0.54,0.81)**	**0.66 (0.54,0.79)**
Peripheral details	0.56 (0.39,0.73)	0.52 (0.34,0.70)	0.46 (0.28,0.64)	0.50 (0.33,0.68)	0.55 (0.39,0.71)	0.51 (0.34,0.67)

Note: Only time intervals from 0-5 s and 5-10 s after stimulus onset are reported. Significant validity coefficients are printed in bold typeface.

Table 5

Areas under the ROC curves (and 95 % confidence intervals) for the eye-blink data.

	Arousal induction		**No arousal induction**		**Across arousal**	
Time after stimulus:	0 – 5 s	5 – 10 s	0 – 5 s	5 – 10 s	0 – 5 s	5 – 10 s
Number of blinks:						
Immediate CIT						
Central details	0.67 (0.48,0.87)	**0.70 (0.52,0.88)**	0.61 (0.41,0.81)	**0.73 (0.55,0.90**)	0.64 (0.47,0.81)	**0.71 (0.56,0.87)**
Peripheral details	0.59 (0.39,0.79)	0.55 (0.35,0.76)	0.51 (0.29,0.73)	0.54 (0.34,0.75)	0.55 (0.36,0.74)	0.55 (0.37,0.73)
Delayed CIT						
Central details	0.68 (0.49,0.87)	**0.77 (0.61,0.93)**	0.51 (0.31,0.71)	**0.72 (0.55,0.90)**	0.59 (0.42,0.75)	**0.75 (0.60,0.89)**
Peripheral details	0.63 (0.44,0.82)	0.53 (0.33,0.73)	0.58 (0.38,0.78)	0.58 (0.38,0.78)	0.61 (0.43,0.78)	0.55 (0.37,0.74)
Across time of test					**Across arousal and time of test**	
Central details	**0.68 (0.51,0.85)**	**0.74 (0.58,0.89)**	0.55 (0.38,0.72)	**0.72 (0.58,0.88)**	0.61 (0.46,0.76)	**0.73 (0.59,0.87)**
Peripheral details	0.61 (0.44,0.78)	0.54 (0.35,0.73)	0.55 (0.36,0.74)	0.56 (0.38,0.74)	0.58 (0.41,0.75)	0.55 (0.38,0.72)

Note: Only time intervals from 0-5 s and 5-10 s after stimulus onset are reported. Significant validity coefficients are printed in bold typeface.

For groups pooled over time, significant validity coefficients were obtained for both fixation measures during the two interesting time intervals after induction of arousal. If no arousal induction occurred, the average fixation duration enabled a valid classification only during stimulus presentation. In contrast, the number of fixations allowed for a valid differentiation only after stimulus offset.

Finally, significant validity coefficients were observed for both fixation measures for the time intervals from 0 to 5 s and from 5 to 10 s, when pooling guilty subjects across both experimental manipulations.

Similar to the fixation data, the number of blinks allowed for a valid group differentiation only for central details. Table 5 shows that validity coefficients were consistently higher than chance level for the time window after stimulus offset (5-10 s). A comparable pattern was found across experimental manipulations of arousal induction and time between mock-crime and CIT. Additionally, when pooling across the time of test, the arousal induction group already showed scores above chance classification during stimulus presentation (0-5 s).

Incremental validity of ocular measures

As can be seen in Table 6, the combined physiological score as well as the combined ocular score allowed for predicting group membership (guilty vs. innocent) in simple logistic regressions although a much larger effect size was observed for the former model. When both measures were included into the same regression model, the autonomic measure became nominally stronger, while the ocular measure became non-significant. This result indicates that the autonomic measure accounted for the predictive variance in the ocular measure, and contained significant incremental variance beyond the ocular measure.

Table 6.

Results of logistic regression models predicting group membership (guilty vs. innocent) from autonomic and ocular measures as well as their combination.

Model		β	p	Nagelkerke's R^2
Model 1: Only autonomic measures				.78
	Intercept	-2.33	< .0001	
	Combined autonomic score	17.01	< .0001	
Model 2: Only ocular measures				.34
	Intercept	-0.24	.35	
	Combined ocular score	-3.57	< .001	
Model 3: Autonomic and ocular measures				.81
	Intercept	-2.84	< .0001	
	Combined autonomic score	21.52	< .0001	
	Combined ocular score	-1.84	.69	

*Note.*Values were determined using a bootstrapping approach with 10,000 samples.

2.4 Discussion

This study aimed at estimating the external validity of the CIT by constructing an experimental setting that resembles certain aspects of field conditions that might be relevant for CIT validity. Especially the influence of emotional arousal on the CIT's validity was investigated. To further address real-life conditions, the amount of time between mock-crime and application of the CIT was manipulated. Even if the relevant details could only be perceived incidentally, explicit memory for central details was 95.3% (*SD* = 9.9%) at the immediate testing occasion and 87.6% (*SD* = 14.8%) after two weeks. Overall, memory for central details was much better in comparison to peripheral ones, and this effect was more pronounced in the arousal induction group.

2.4.1 Autonomic responses

The physiological responses during the CIT supported the findings for the explicit memory performance by revealing an advantage of central over peripheral details, and stronger responses to relevant as compared to irrelevant central CIT items in all autonomic response measures. A diminished standard difference score at the delayed measurement occasion was observed for most experimental groups and all measures except for the electrodermal response. For the heart rate data, an increased temporal stability for central as compared to peripheral details was found as trend. This result was not statistically significant and should therefore be interpreted with caution, but it delivered some additional evidence for a stable differentiation between relevant and irrelevant items over time depending on the physiological measure used and the type of critical detail.

There were no significant effects of emotional arousal on the physiological responses during the CIT, but for respiration and heart rate measures trends were observed for temporally more stable responses in the arousal induction group. Prior research reported that RLL validity might be higher in the field compared to laboratory settings (Elaad et al., 1992). Specifically, a reanalysis of the CIT data of 30 Japanese suspects later found guilty of serious crimes indicated that respiratory apneaic responses were more pronounced within CIT questions directly related to the execution of the crime (and hence perhaps involve the retrieval of emotionally arousing facets of the crime) as compared to more peripheral details that could have also been encoded in a less arousing situation before or after the crime (Suzuki, Nakayama, & Furedy, 2004). The conventionally used skin resistance response measured by Suzuki and colleagues (2004) did not show the same specific sensitivity which led to the assumption that respiratory suppression might be more sensitive to emotionally arousing crime facets than electrodermal measures. Further support for sensitivity differences in physiological measures comes from studies reporting a differential habituation profile across the CIT examination. For the respiratory (e.g., Ben-Shakhar & Elaad, 2002) as well as for the heart rate responses (Gamer, Gödert, et al., 2008), it was shown that these measures are temporally more stable and more robust against experimental manipulations than electrodermal responses, indicating that these measures reflect partly different psychological processes. This interpretation also fits with the observation that countermeasures are less effective on respiratory as compared to electrodermal responses (e.g., Ben-Shakhar & Dolev, 1996). It remains an important challenge for future research to further examine the response specificity that could explain the incremental validity of different measures and might substantiate theoretical concepts explaining physiological responding in the CIT (Ben-Shakhar & Furedy, 1990; Verschuere et al., 2004; Verschuere & Ben-Shakhar, 2011).

An analysis of the combination of electrodermal, respiratory and heart rate responses according to a previously described classification function (Gamer, Verschuere et al., 2008) revealed similar effects as found for the single measures. Thus, advantages for central over peripheral details and for the immediate over the delayed testing session were observed. The calculated validity coefficients demonstrated that central details are more suitable for the detection of crime relevant information. The observed areas under the ROC curve ranging from .90 to 1.00 are in line with previous studies (Gamer, Verschuere, et al., 2008) and indicate that the CIT is well suited to detect concealed crime-related knowledge even when this information was only incidentally encoded. In line with prior research, peripheral details were less well recognized during the CIT and resulted in a higher risk for false negative classification especially during a more realistic application with a delayed CIT (Carmel et al., 2003; Gamer et al., 2010; Nahari & Ben-Shakhar, 2011).

This is the first study that systematically varied the amount of emotional arousal during information encoding and importantly a positive influence of arousal on the CIT validity was found. Specifically, validity coefficients for central details declined as a function of time in the group without arousal induction, but remained temporally stable when participants experienced enhanced arousal during the crime. Moreover, even though participants in the arousal induction group were not able to recognize as much peripheral details as examinees who were not disturbed during the mock-crime, validity coefficients were comparable between both groups. Thus, it seems possible that emotional arousal during the crime enriched memory for relevant crime details which led to more stable physiological response differences in the subsequently accomplished CIT. These effects were more pronounced at the delayed testing occasion, presumably due to enhanced memory consolidation between the mock-crime and the CIT (e.g., Sharot & Yonelinas, 2008).

2.4.2 Eye-movements

Furthermore, this study aimed at examining whether a valid detection of crime-related memory is possible on the basis of ocular measures in a standard CIT setting. So far, heterogeneous results were published regarding this topic. Interestingly, eye movements were suitable for detecting concealed information by showing a reduced number of fixations and prolonged fixation durations for relevant compared to neutral CIT items. Comparable to the explicit recognition rates, this effect was restricted to central crime details and resembles eye movement based memory effects obtained in basic research (e.g., Althoff & Cohen, 1999; Hannula et al., 2010). Thus, guilty subjects showed fewer fixations when viewing pictures of relevant details due to their prior contact with them during the mock-crime. This effect was most prominent during picture presentation and decreased after picture offset. However, even during the 5 s immediately following stimulus presentation, the number of fixations in the area of the screen where the stimulus was previously shown was reduced. A similar pattern was found for the average fixation duration, revealing longer fixations on central relevant details during stimulus presentation and after stimulus offset. Both fixation measures enabled a valid differentiation between guilty and innocent participants during the first and second time interval, when pooling data across experimental manipulations. The number of fixations and the average fixation duration were negatively correlated, indicating that fewer fixations were accompanied by enhanced fixation duration. This pattern indicates that persons with crime-related knowledge show less exploration of crime-related stimuli instead of avoiding to look at them.

In line with prior research (Fukuda, 2001; Leal & Vrij, 2010) the eye-blink data enabled detection of concealed information. A reduction in eye-blinks occurred after presentation of pictures showing central details of the mock-crime and this effect was most prominent after stimulus offset (5-10 s after onset of stimulus presentation) while the participants saw a uniformly gray screen. The same time interval was crucial for a differentiation between guilty and innocent subjects, as shown by the calculated validity coefficients. When pooling data across the immediate and delayed test, a valid classification was also possible for the number of eye-blinks during stimulus presentation. The late onset of decreased eye-blinks after presentation of crime-related information in the current study contrasts with Fukuda (2001), who reported a reduction of blinks at the end of stimulus presentation but a rebound effect after stimulus offset. However, since issues of multiple testing were not taken into account by this author, it is unclear how reliable this rebound effect really is. Leal and Vrij (2010) discuss that cognitive load might be a viable theoretical explanation for the reduced blink rate on relevant CIT details. Following this line of reasoning, the presentation of CIT items might initially elicit an orienting response (that is larger for recognized details), which initiates a more controlled subsequent processing. The former process seems to be more related to the pattern of physiological responses that is usually observed in the CIT (Gamer, 2011a), whereas the latter process might be sensitive to cognitive load, which is potentially reflected in the reduced blink rate several seconds after stimulus onset. Moreover, since correlations between eye-blinks and fixation measures were only moderate in the current study, it seems possible that different processes underlie this response pattern. Alternatively, inhibition instead of cognitive load may account for the reduced amount of eye-blinks. In fact, it was shown that startle eye-blink responding to crime-related pictures was affected by inhibition but not orienting (Verschuere, Crombez, et al., 2007). In line with this data, our finding of reduced eye-blinks after picture offset might indicate a prolonged behavioral inhibition, which is potentially linked to attempts of the examinees to monitor or control bodily responses to effectively conceal crime-related knowledge. Future studies using explicit manipulations of cognitive load and inhibition might elucidate their influence on the temporal pattern of eye-blinks in more detail.

Even though physiological responses and ocular measures separately allowed for a valid detection of crime-related knowledge, they were only partly correlated. Moderate associations were found between the combined physiological score and the number of fixations during stimulus presentation and 5 s afterwards. Thus, stronger physiological responses to relevant details were accompanied by larger decreases in the number of fixations and longer fixation durations. No association with the eye-blinks and the combined physiological score was found. This differential pattern of correlations further supports the idea that fixations might be more related to an initial orienting response whereas eye-blinks reflect subsequent processes related to cognitive load or inhibition.

Although, no specific hypothesis regarding an effect of the emotional arousal and time between mock-crime and CIT manipulations on ocular measures existed, an exploratory analysis was conducted to test for an influence of these experimental manipulations. Interestingly, some evidence was found for an improved detection of concealed information during a delayed CIT after arousal induction. Specifically, the number of fixations revealed an advantage for central compared to peripheral details during the delayed CIT after the induction of arousal. In contrast, the group without arousal induction did not show this effect. Additionally, the data on eye-blinks showed a stronger response in the arousal induction group compared to the no arousal induction group during the delayed CIT. No differences between these two groups were found for the immediate test. Overall, for some ocular measures emotional arousal seems to enhance response differences between relevant and irrelevant CIT details during a delayed test. However, differences for ocular measures were small and partly depended on the type of critical detail as well as the time window. Furthermore, no effect of the experimental manipulations was found for the fixation duration.

Thus, these effects have to be replicated by future studies to ensure that the observed group differences are reliable.

Regarding validity coefficients, the areas under the ROC curves for the obtained ocular measures were only moderate and did not exceed a value of 0.82. Especially the number of fixations and the average fixation duration revealed differences in valid classification based on single groups compared to groups pooled over experimental manipulations. These differences are presumably driven by the small group and moderate effect sizes, respectively. The results for the eye-blinks enabled a valid differentiation between guilty and innocent subjects after stimulus offset for each guilty group and for groups pooled over experimental manipulations. However, validity coefficients for all ocular measures were substantially smaller as compared to other physiological measures. For example, electrodermal responses alone typically reach validity coefficients between 0.80 and 0.95 (Ben-Shakhar & Elaad, 2003). In the current sample, ocular measures alone were predictive of group membership (guilty vs. innocent) but this variance was also accounted for by the combined physiological score. Thus, ocular measures did not add incremental validity to traditional autonomic response measures. This could be a hint that autonomic and ocular response changes are based on similar psychological processes during the CIT examination. Importantly, the current study was not specifically designed for examining the incremental validity of single measures and therefore a bootstrapping approach was used for predefined combinations of autonomic and ocular responses, respectively. The huge difference in sample size between guilty and innocent subjects, as well as the small number of innocents compared to the large number of potential predictors (4 autonomic and 3 ocular measures) precludes an application of classification approaches using single measures. Therefore, the examination of larger samples will be important for future studies in this domain.

2.4.3 Conclusion and outlook

Taken together, the current findings indicate high CIT validity under realistic and thus crucial experimental conditions, including emotional arousal and delayed CIT application. This result is contrary to other realistic mock-crime studies that reported decreased detection rates under realistic test conditions (e.g., Nahari & Ben-Shakhar, 2011). However, as predicted from results in basic research on emotional memory (e.g., Christianson, 1992), in the current study the arousing event of a mock-crime seemed to increase memory for central details of the scenario at expense of memory for more peripheral details. Thus, even if the CIT mainly relies on memory processes, emotional factors seem to modulate the encoding and receiving of relevant crime details.

In addition, the current study demonstrates that the conjunction of basic research on eye movements as an indirect memory test with forensic applications is a fruitful approach that could further advance the knowledge on theoretical and applied issues surrounding the CIT. Although the validity of ocular measures alone is only moderate, the main advantage of eye-tracking is the possibility to measure these responses unobtrusively. Such application seems highly attractive for certain forensic issues but prompts ethical concerns that have not been comprehensively addressed so far. Future studies should also focus on the effect of countermeasures, as eye movements and blinks are at least partly under voluntary control and it is currently unclear to what degree guilty subjects might deliberately alter their pattern of ocular responses to appear innocent.

In sum, even though more research is needed to clarify the dependence of single physiological measures on situational factors such as emotional arousal, the current data support the application of the CIT in forensic cases.

3. Study II

3.1 Introduction

The second study of the current thesis manipulated the encoding context for relevant details in guilty and innocent subjects and examined the neural responses in a CIT that asked for these details. Differences in encoding context were examined in some CIT studies. For example, a valid detection of knowledge in persons who only intended to commit a mock-crime were reported based on electrodermal responses (Meijer et al, 2010) or RTs (Noordraven & Verschuere, 2013). Heterogeneous results were reported regarding the detection of knowledge in guilty subjects and informed innocents (Ben-Shakhar et al., 1999; Bradley & Rettinger, 1992; Bradley & Warfield, 1984; Gamer, 2010). Recently, some studies used fMRI in combination with CIT designs and aimed to deliver new insights in the mechanisms underlying the CIT (for a review see Gamer, 2011b). These studies revealed that critical items elicited higher activity compared to neutral alternatives in a ventral fronto-parietal network, consisting of the bilateral inferior frontal gyrus (IFG), right middle frontal gyrus (rMFG) and right temporoparietal junction (rTPJ) (Gamer, 2011b).

Most of these previous fMRI-based CIT studies used the playing card paradigm and asked their subjects to hide knowledge about a specific card. Thus, to overcome limitations of previous CIT studies using very artificial and highly restricted laboratory designs, a complex, realistic mock-crime scenario was developed for the current study to test the fMRI-based CIT approach under rather realistic conditions. Moreover, it remains an open question whether neural responses during the CIT are more sensitive for differences in information processing compared to autonomic measures. To manipulate the encoding context, three different experimental conditions were generated: Some persons only intended (PLAN group) to commit a mock-crime, while some persons enacted the mock-crime (ACT group) in reality. An additional group of informed subjects (INNOCENT group) was tested with the same CIT.

These subjects enacted a neutral task during which they acquired knowledge of some relevant details. As the PLAN group and the INNOCENT group only knew half of the details, these subjects were uninformed innocents regarding the other half of the relevant information. Thus, the current design enables comparisons between groups that differ regarding the context of memory encoding for the same relevant details as well as traditional comparisons between guilty and innocent subjects.

For all groups increased activation in the previously reported ventral fronto-parietal network (Gamer, 2011b) was hypothesized for the comparison between known relevant details and irrelevant details. In addition, differential changes in BOLD activity in the ACT group compared to the PLAN group or the INNOCENT group were explored. Increased emotional arousal was expected for subjects who committed the mock-crime (ACT group) compared to the persons who only planned the mock-crime (PLAN group) or who encoded the relevant details during an errand (INNOCENT group). In detail, for the ACT and PLAN group activation differences in the supramarginal gyrus (SMG) were assumed, as prior research revealed that enactment influences the hemodynamic response in this region during memory tasks (Russ, Mack, Grama, Lanfermann, & Knopf, 2003). Differences in amygdale and hippocampus activations were predicted between the ACT and INNOCENT group, as basic research repeatedly found stronger activation for the retrieval of emotional compared to neutral information in these regions (for a review see Costafreda, Brammer, David, & Fu, 2008).

3.2 Methods

3.2.1 Participants

This study was approved by the local ethics committee and conducted according to the principles expressed in the Declaration of Helsinki. All participants gave written informed consent and were paid 50 Euro for participation. The final sample consisted of 60 righthanded participants (18 women) with a mean age of 25.3 years (*SD* = 3.1). Most persons were students from various faculties (70%). During data collection, eight persons had to be excluded from the study for several reasons and were replaced by new subjects. Reasons for exclusion were technical difficulties (*n* = 3), an incomplete fulfillment of the mock-crime (*n* = 3) or alcohol intoxication (*n* = 2).

3.2.2 Design

A realistic mock-crime procedure was either only planned (PLAN) or really enacted (ACT) by subjects in the guilty groups. Persons who belonged to the PLAN condition knew only relevant details that belonged to the planning phase, whereat the ACT group knew all relevant details. In addition, a third group of persons fulfilled a non-criminal task before they were examined with the same CIT (INNOCENT). These subjects knew half of the relevant details from a neutral context, and found out during the CIT that these details were also part of a mock-crime. Each group consisted of 20 subjects and participants were randomly assigned to their respective experimental condition by means of a predefined list.

3.2.3 Procedure

The experiment was conducted in three different stages: a planning phase (day one to seven), an action phase (day seven), and a CIT examination (day seven or day eight).

Stage 1. Participants arrived at the laboratory individually and were met by an assistant who obtained written informed consent and explained the experiment in further detail. Guilty and innocent participants were informed that a person would contact them during the upcoming week via email and provide further details regarding their respective task. They should respond to these emails as fast as possible and answer questions presented to them. Furthermore, they were asked to come back to the department one week later to fulfill their respective task and the fMRI-based CIT would take place with one day delay.

Guilty participants of both conditions (PLAN and ACT), received emails on day one, day three and day five after they signed informed consent. These emails contained twelve details that would be used as relevant items in the later CIT examination. This was unknown by the participants. The respective details are printed in italics below. In the first email, the contact person, named *Otto*, told the subjects that he works in the *Institute of Applied Nuclear Research* and wants to steal a *CD* with important *study information* from a *colleague*. Participants were asked to check the map of the University Medical Center Hamburg-*Eppendorf* (UKE) and look for the building called *Campus Lehre*. This was the building where the mock theft was about to take place and it was shown in *blue color* on the official campus map that participants had to retrieve to report the building's number to Otto by email. This interaction was included to ensure that subjects actively prepared for their task. The second email contained further details regarding the theft. Participants learned that a key would be hidden for them in a *Lavazza can* and that they should use that key to open the locker number *220* located in the building’s basement. Participants were instructed to open a *textile bag* inside the locker, and take the CD and all money from the wallet. This time participants had to check the homepage of the UKE to look up *when the lockers are available* and report this to Otto again. The third and last email informed the participants to fulfill the theft in no longer than 15 minutes, otherwise there would be a high risk for being catched. Additionally, the participants were asked to check the CD's content at home and send an email with the number and type of files to Otto.

Innocent participants received their emails on the same days as guilty subjects, but they had to plan an errand instead of a mock-theft. The emails provided innocent subjects with knowledge about six details that were also involved in the mock-crime scenario. They were asked by *Otto* to meet a student assistant at the building *Campus Lehre* to get a *CD* with important *study information*. In parallel to the guilty subjects, the innocent persons had to check the University Medical Center Hamburg-*Eppendorf* campus map in the internet and thereby became aware of the *blue coloring* of the building's location. After completion of the errand, participants should check the CD and inform Otto about the number and type of files on it.

Stage 2. After successful completion of the planning phase, half of the guilty subjects committed the mock-crime (ACT group). During the mock-crime, these participants could encode twelve additional details that were later used in the CIT. These relevant details are printed in italics below. Guilty ACT subjects could see that the key knob was a *car* and that *Fackelmann* was written on the locker's lock. Furthermore, they stole a *20 Euro note,* and saw a picture of the movie *Cheyenne* pinned on the lockers inside. Next to the textile bag, a *hat* and a *bottle of ice tea* were placed inside the locker. On their way, subjects had to pass a *snack automate* and a *bookstore*. A sticker showing the word *Sicherung* was placed on the cover of the CD they had to steal. *Four* files were saved on the CD, including one picture showing a *blue curvest* and three text files entitled *Randomization*.

The second half of the guilty subjects (PLAN group) were unexpectedly stopped before they could commit the mock-crime. These subjects only knew details regarding the planning phase, and went straight to the MRI scanner to undergo the CIT examination.

Innocent participants met the student assistant to fulfill their errand one week after they signed the informed consent. They got the chance to encode six details that were also presented to subjects in the guilty ACT group, but in a non-criminal context. They passed a *bookstore* and a *snack automate* before they got the *CD* with a sticker showing *Sicherung* on the cover. They informed Otto that *four* files were saved on the CD. The CDs content was identical to the ACT group, one picture showing a *blue curvest* and different text files entitled *Randomization*. The student assistant was unknown to subjects in the INNOCENT group and did not talk to them for a longer time to assure standardized conditions. The same student assistant was responsible for replacement of all relevant items after subjects in the ACT group fulfilled the mock-crime.

Stage 3. The CIT was conducted by an examiner unknown to the subjects, who was uninformed about the respective participant's group condition. Subjects were instructed not to mention their respective group condition until the whole experiment was completed. Additionally, participants were told that the experiment was designed to check whether they could cope with the polygraph test and convince the examiner that they were innocent. An additional reward of 30 Euro was offered for a successful performance in the CIT to increase their motivation. First, a practice session outside the scanner was completed to familiarize participants with the CIT procedure. Two questions, unrelated to the mock-crime, were presented and each question was followed by five answer options, presented either as words or as pictures. Subjects were instructed to always press the right button after stimulus appearance. Only when a small red circle was visible next to the stimulus they should press the left button. This design resembles the three item CIT approach that was frequently used in previous CIT studies (cf. Gamer, 2011b). After successful completion of the practice task, participants were placed in the scanner. Button presses had to be given by the index and middle finger of the participants’ right hand.

The CIT consisted of 24 questions about the mock-crime, twelve about details from the planning phase and twelve about details from the enactment phase. A complete list of questions and items is presented in Table 7.

The order of relevant and irrelevant items following a question were pseudorandomized and remained constant across participants. Three blocks of 24 questions were presented. The first block was always identical for all participants. Blocks two and three included the same questions, with different orders of questions and mixed orders of the respective answer options. Six different sets for questions and answer items were prepared in a predefined list and subject numbers were randomly assigned to specific blocks.

Table 7.

Questions and details used during the CIT.

Questions PLAN-phase	Buffer item	Irrelevant item	Irrelevant item	Target item	Relevant item
What is the name of the district you went to?	Neustadt	Altona	Harburg	Hohe Luft	Eppendorf
Where did your contact person work?	International Museum of maritime Stuff	Max-Planck Institute for Meterology	Controlling and quality research institute	Institute of tropical research	Institute of Applied Nuclear Research
What kind of information was your contact person interested in?	Address book	Statement of account	Pictures security camera	Password	Study information
What was the name of the building you visited?	Bookhall	Agency	Bank institute	Post	Campus Lehre
What was the colour of your goal area on the map?	Pink	Orange	Brown	Turquoise	Blue
Where was the key hidden that you used?	Blue box	New York box	Plastic box	Wooden box	Lavazza box
What is the name of your contact person?	Egon	Mario	Jens	Klaus	Otto
Who owns the key that you used?	Wife	Friend	Neighbour	Mother	Colleague
What was the number on the locker?	22	6	134	217	220
What did you steal from the locker?	I-pod	Notebook	Mobile phone	Arm watch	CD
What time are the lockers available?	9h	7h	8h	7.30h	6.30h
What contained the things you have stolen?	Suitcase	Bag	Rucksack	Plastic bag	Textile bag

Questions ACT-Phase	Buffer item	Irrelevant item	Irrelevant item	Target item	Relevant item
What kind of key fob was attached to the key?	Ribbon	Knife	Cord	Rose	Car
What did the locker look like that you opened?	Purple coloured	Black coloured	Entitled Nautic	Extra long shape	Entitled Fackelmann
What did you see on the flyer inside of the locker?	Art studio	ZEIT journal	Periodic system	Postcard Sea	Cheyenne movie
How much money did you steal from the locker?	100 Euro	10 Euro	5 Euro	50 Euro	20 Euro
What piece of clothing did you see in the locker?	Scarf	Running shoes	Base cap	Gloves	Hat
What did you see in addition in the locker?	Deodorant	Snack bar	Folder	Book	Bottle of ice tea
What did you pass to go to the stairs?	Wardrobe	Plant	Post office	Parking automate	Snack automate
What shop was located in the building?	Perfumery	Grocery store	Flowers store	Starbuck	Bookstore
Within the stolen item, what infomration did you see?	Volcanplot	Bars	Boxplot	Curvest	Westernplot
Within the stolen item, how many files were there?	Six	Three	Seven	Five	Four
Within the stolen item, what was the content of the files?	Application for leave	Report	Bill	Diploma thesis	Randomization
What was written on the sticker that was placed on the stolen item?	Top Secret	Reorder	Proof	Copy	Backup

Note. Questions and items are loosely translated in the present table. Half of the described items were presented as pictures, the other half as words.

During the CIT, each question was followed by five answer options. For trial timing see Figure 9A. The first item was always a buffer and not included in the analysis. The remaining four items included two irrelevant items, one relevant item that mentioned or depicted the correct answer, and one target item that required a different button press. The targets were identical for all participants and always included a small red circle next to the presented word or picture. Two examples of CIT questions are shown in Figure 9B. In contrast to previous studies (e.g., Gamer et al., 2007), we did not use a set of previously memorized specific target items but defined a perceptual feature (i.e., a small red circle) to mark the targets. This was done because it seemed too difficult to encode a large number of target items which would have been necessary because of the 24 CIT questions that were used in the current study. Reaction times and button presses were monitored during the CIT and subjects who did not respond, responded incorrectly or too slowly were reminded to pay attention to the task during the short breaks between blocks.

After subjects completed the CIT in the scanner, they were asked to reveal their respective group condition and to return things they kept from their respective task (money and/or CD). In addition, all subjects retrospectively rated their subjective nervousness during the respective task (i.e., planning a mock-crime, committing a mock-crime, committing an errand) on a Likert scale ranging from 1 *(not nervous at all)* to 5 *(nervous all the time)*.
Afterwards, an explicit memory test was conducted to check which relevant details the participants remembered correctly. During this test, all CIT questions were presented in a multiple-choice format on a laptop and subjects had to choose the correct answer option. If the correct answer was unknown to the subject, which was the case for half of the questions in the PLAN and the INNOCENT group, participants were asked to guess the correct answer.

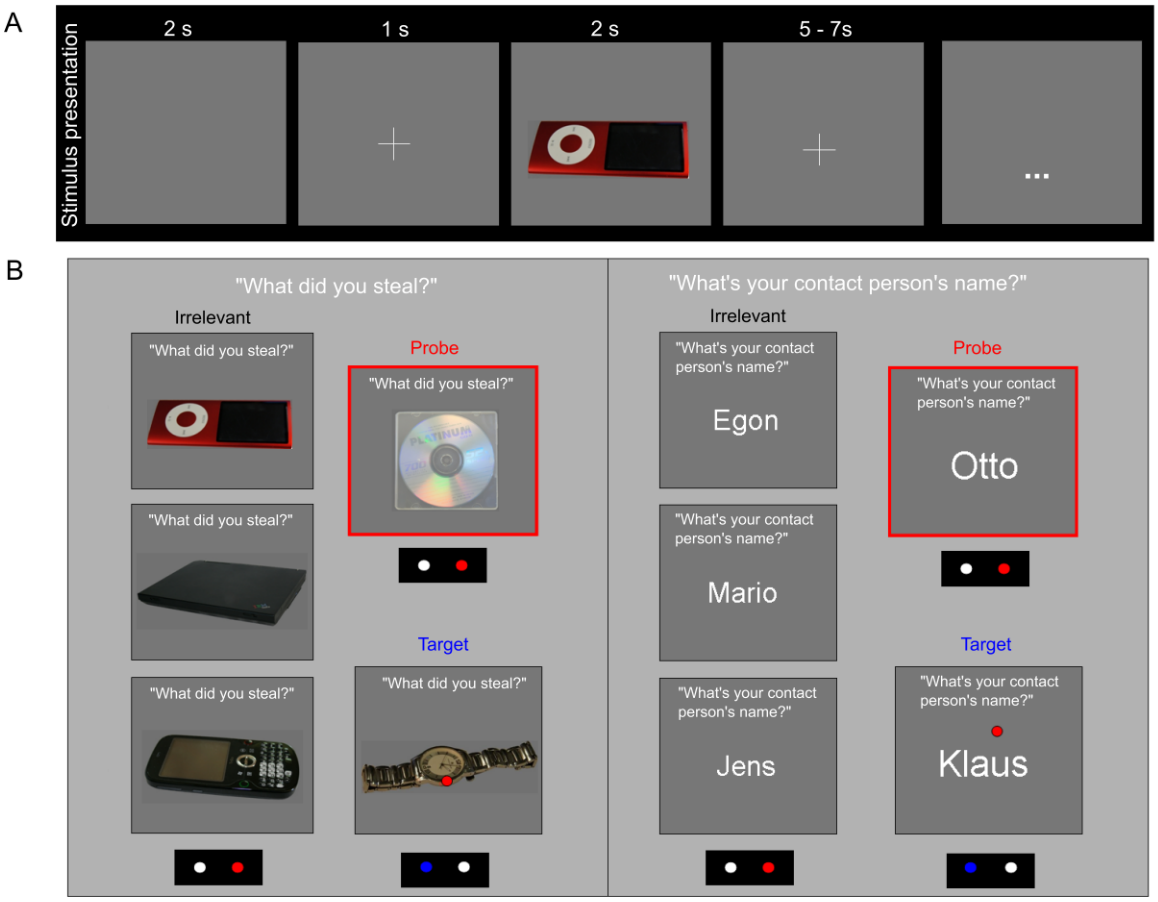


Figure 9. A) Event sequence during the CIT (Study II). B) Examples for item types and corresponding button presses.

To explore long-term memory for the relevant details, subjects were invited to accomplish the same explicit memory test again six month after the CIT using an online internet questionnaire. From the original sample, 42 subjects completed the questionnaire. Of these belonged n = 12 to the PLAN group, n = 15 to the ACT group and n = 15 to the INNOCENT group.

3.2.4 Data acquisition

Presentation of CIT questions and items as well as recording of behavioral responses was performed with Presentation software (Neurobehavioral Systems, Albany, CA, USA). Functional MRI was performed on a 3 T MR whole body scanner (Magnetom Trio, Siemens) using a 32-channel head coil. Forty continuous axial slices (slice thickness 2 mm, 1 mm gap) were acquired in each volume using a T2*-sensitive gradient echoplanar imaging (EPI) sequence (TR: 2390 ms; TE: 25 ms; flip angle: 80°; field of view: 216×216 mm; voxel size: 2×2×2 mm^3, GRAPPA with PAT-factor 2). Isotropic high-resolution (1×1×1 mm^3) structural images were acquired using a T1-weighted coronal-oriented MPRAGE sequence with 240 slices. Skin conductance was recorded using a constant voltage system (0.5 V, Biopac MP100 System, Biopac. Inc) with electrodes placed on the thenar and hypothenar eminences of the participant's left hand.

3.2.5 Data processing and analysis

Subjective nervousness ratings All subjects rated the subjective nervousness during their task (i.e., planning a mock-crime, committing a mock-crime, fulfilling an errand) on a Likert scale ranging from 1 *(not nervous at all)* to 5 *(nervous all the time)* after the CIT. Six subjects in the PLAN group did not answer this question, therefore this group includes only 14 subjects for this analysis. A one-way ANOVA was calculated on the average nervousness scores with the between-subjects factor group.

Recognition task. From the explicit memory test, the sum of correctly remembered items was calculated for each subject. The ACT group could recognize a maximum of 24 relevant details, while the PLAN group could only know twelve relevant details from the planning phase. The INNOCENT group was aware of six relevant details from the planning phase and six relevant details from their errand task. The recognition rate was calculated based on all items that could be known to the respective group to enable comparison between groups. The same measure was calculated for the 42 persons who accomplished the explicit memory test six month later.

Skin conductance responses. Amplitudes of skin conductance responses that began between 1 and 3 s after stimulus onset were scored as stimulus-evoked responses and measured as changes in micro-Siemens if they exceeded a threshold of 0.01 μS. The amplitudes were log-transformed according to the formula provided by Venables and Christie

(1980). Trials with missing behavioral responses (0.15% of all trials) as well as buffer and target items were excluded from further SCR analyses. To facilitate a comparison between examinees, standard difference scores were calculated (cf. Gamer et al., 2010). Within each participant the difference between the relevant item and the mean of the irrelevant items within each question was calculated and divided by the standard deviation of all irrelevant items of the CIT. Subsequently, the response differences for each question were calculated and scaled with the individual response distribution of the irrelevant items within each subject. These scores were averaged across all CIT questions (ACT group) or averaged separately for CIT questions where the correct item was known or unknown, respectively (PLAN and INNOCENT group).

Imaging data. Image processing and statistical analyses were carried out using Statistical Parametric Mapping (SPM8; Wellcome Department of Imaging Neuroscience, London, UK) running under Matlab R2009b (Mathworks, Inc., Natwick, MA, USA). The first four volumes of each time series were discarded because of T1 equilibration effects. Volumes were slice time corrected and realigned to the mean EPI to correct for movement artifacts. Subsequently the structural T1 image was coregistered to the mean EPI image. T1 images were segmented and resulting transformation parameters were used for spatial normalization of EPI and T1 images to MNI space. Finally, data were smoothed with a 6 mm full-width at half maximum (FWHM) isotropic Gaussian kernel.

Data analysis was performed using a general linear model approach. The first-level design matrix of subject's belonging to the PLAN group or the INNOCENT group consisted of 4 x 2 regressors. For these groups, one regressor was created for each item type (relevant, irrelevant, buffer and target) for either the known or the unknown condition. Events were modeled as stick functions for each stimulus onset. For the ACT group only four regressors, representing the item types were included in the design matrix, because these subjects were aware of all relevant details. For all groups contrast images were defined for known relevant details minus irrelevant details. In addition, for the PLAN and INNOCENT group the contrast images for unknown relevant minus irrelevant details were defined. On the second level, a flexible factorial design matrix was constructed with six regressors. For each group, one regressor for the known relevant details and one regressor for the respective irrelevant details were included. This model enabled exploration of the interesting contrast (known relevant minus irrelevant details) in each group. However, it is important to note, that knowledge of the ACT group was based on 24 relevant details, while the PLAN and INNOCENT group only knew half of these relevant details.

To enable comparison between groups (ACT versus PLAN; ACT versus INNOCENT) under identical conditions, two additional first level models were constructed. In each model, the relevant details of the ACT group were split according to whether they were known to the PLAN or INNOCENT group, respectively. Thus for the ACT group the amount of known details was reduced based on the knowledge of one of the other groups. Contrast images were defined for the known relevant minus irrelevant details and the unknown relevant minus irrelevant details. Based on these contrasts, two separate second level models with four regressors each were defined to compare the ACT group with the PLAN or INNOCENT group, respectively. In these models, regressors for the contrast images in the known and unknown condition were defined for each group.

Regions of interest (ROI) were defined based on the coordinates reported in a meta-analysis of fMRI studies on the CIT (Gamer, 2011b). The respective ROIs were the left IFG (x = -44, y = 19, z = 1), the right IFG (x = 39, y = 23, z = -10), the right TPJ (x = 60, y= -48, z = 39) and the right MFG (x = 35, y= 44, z = 23). To correct for multiple comparisons, small volume corrections were applied for each ROI in each group using 12mm spheres. The relative difference between known relevant details and irrelevant details was examined by post hoc analyses on the contrast estimate change averaged across a 5mm sphere around the defined ROI centers. These values were obtained using the SPM-toolbox rfxplot (Gläscher, 2009). To explore additional regions that might differ between the respective item types, a whole brain analysis with family-wise error (FWE) corrected *p*-values ($p_{FWE} < .05$) was conducted within each group for the contrast known relevant minus irrelevant details. For comparison between the ACT and PLAN group and the ACT and INNOCENT group, respectively, small volume corrections in the hypothesized ROIs (i.e., SMG, amygdale, hippocampus) were conducted. For each ROI a mask was generated using the WFU Pickatlas (Maldjian, Laurienti, Burdette, & Kraft, 2003; Tzourio-Mazoyer, et al., 2002). All activations are reported using x, y, z coordinates in MNI standard space.

Statistical analysis. For all analyses the a priori significance level was set to $p = .05$, but marginally significant results are reported at a threshold of $p < .10$. For the fMRI data, results corrected for multiple comparisons are reported with a FWE corrected *p*-value ($p_{FWE} < .05$). Cohens *d* (Cohen, 1988, p.19ff) and Cohens *f* (Cohen, 1988, p. 273ff) are reported as effect size estimates.

3.3 Results

3.3.1 Subjective nervousness

To compare subjective nervousness ratings for the respective task (i.e., planning a mock-crime, committing a mock-crime, fulfilling an errand), a one-way ANOVA with the between-subjects factor group was calculated. This analysis revealed a significant main effect for group ($F(1,52) = 23.336$, $p < .001$, $f = 0.62$), showing the highest score for the ACT group (mean = 2.40; $SE = 0.24$), followed by the INNOCENT group (mean = 1.70, $SE = 0.15$) and the PLAN group (mean = 0.93, $SE = 0.19$). Thus, although subjective nervousness was relatively low in all groups, it was higher in the ACT group and slightly elevated in the INNOCENT compared the PLAN group.

3.3.2 Explicit memory

To compare the percentage of correctly remembered relevant details between the three groups in the immediate memory test, a one-way ANOVA with the between-subjects factor group (ACT, PLAN, INNOCENT) was calculated. As shown in Figure 10A, no difference in memory performance was found between the three groups, $F(2,57) < 1$. For the delayed memory test after six months a 3 x 2 ANOVA was calculated with the within-subjects factor time of measurement (immediate, delayed) and the between-subjects factor group (PLAN, ACT, INNOCENT). Figure 10B reveals a main effect of time of measurement, $F(1,39) = 7.22$, $p = .011$, $f = 0.50$. No further group differences or interaction effects reached statistical significance.

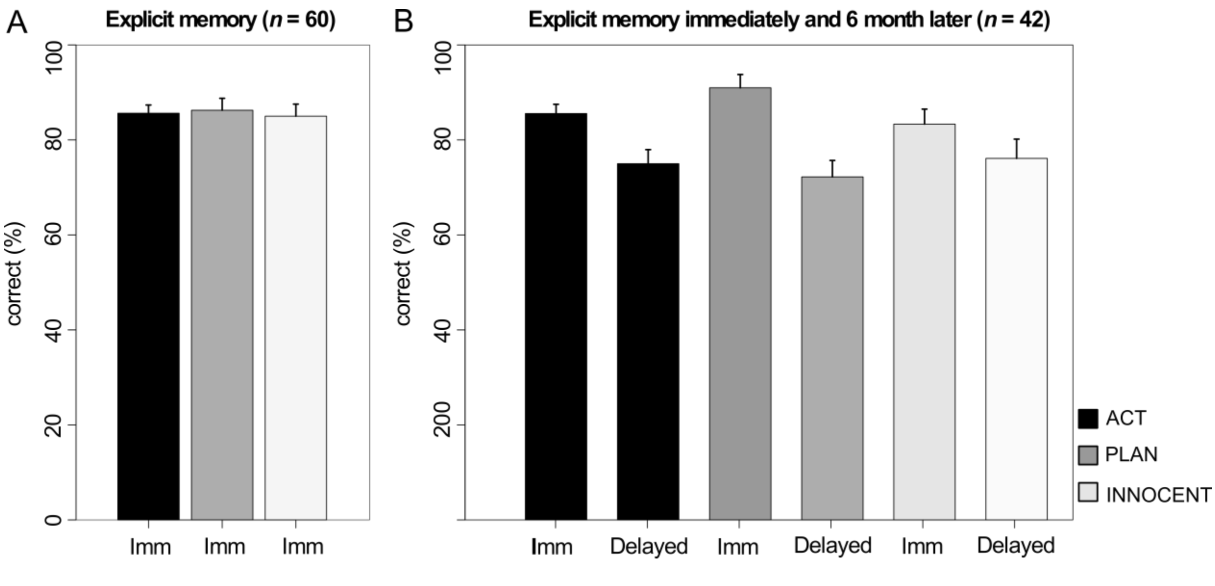


Figure 10. Explicit memory performance (Study II). A) At the immediate (IMM) measurement occasion and B) six month (Delayed) after study participation. All values show the percentage of correctly remembered relevant details for those items that could be known by the respective participants. Error bars indicate standard error of the mean.

3.3.3 Skin conductance

A one-way ANOVA for the difference scores between known relevant and irrelevant details, with the between-subjects factor group (ACT, PLAN, INNOCENT), revealed a significant main effect of the intercept, $F(1,57) = 6.105$, $p = .016$, $f = 0.33$, and a trend for a main effect of group, $F(2,57) = 2.495$, $p < .091$, $f = 0.30$. All other effects were not statistically significant. A one-way ANOVA for the difference scores of unknown details, including the between-subjects factor group (PLAN, INNOCENT) was calculated and revealed no significant results. As shown in Figure 11, only for details known by the participants an increase in skin conductance responses was observed for relevant compared to irrelevant details, and this effect seemed to be slightly more pronounced in the innocent compared to the two guilty groups. If the relevant details were unknown by the participants no such effect was found.

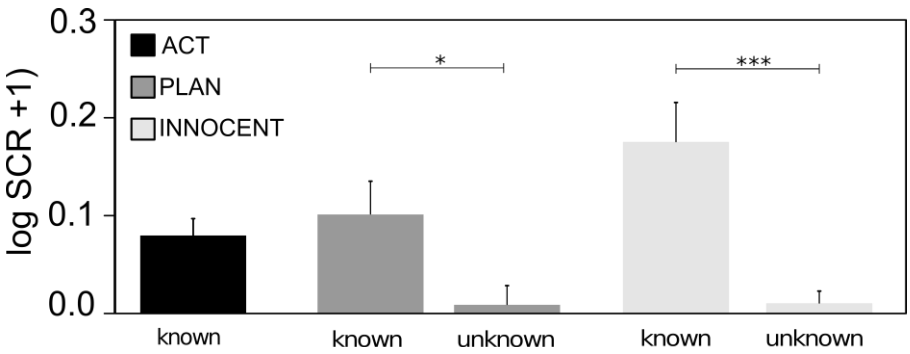


Figure 11. Skin conductance difference scores for relevant and irrelevant details (Study II). Significance is depicted based on *t*-test comparisons between the known and unknown condition, with *** $p < .001$, * $p < .05$. Error bars indicate standard error of the mean.

3.3.4 Imaging data

For each group, significant differences in neural activation between relevant and irrelevant items were observed in the predicted ROIs, more precisely in the bilateral IFG and the right TPJ (see lower part of Table 8 and Figure 12).

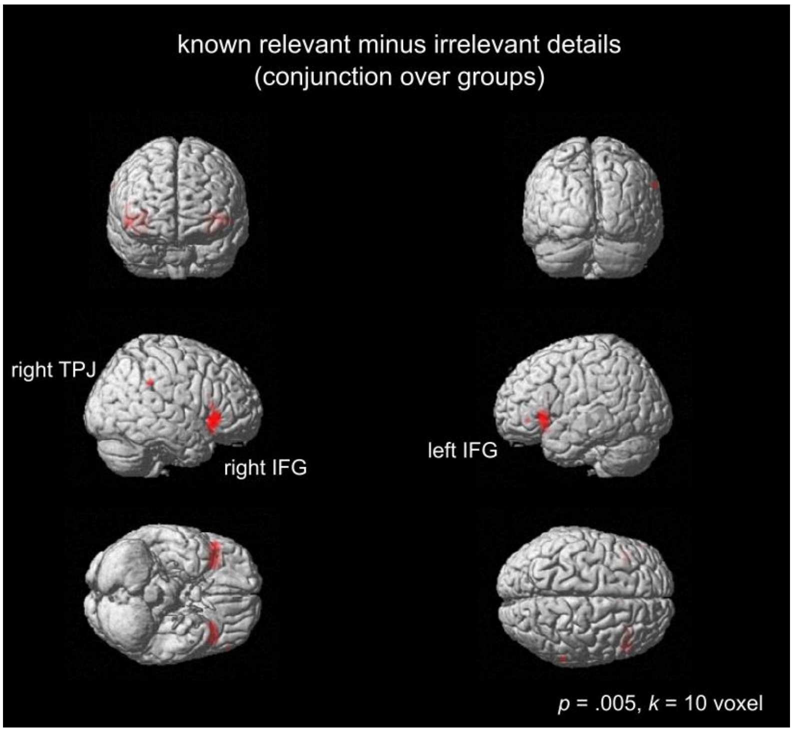


Figure 12. Conjunction analysis across groups (ACT, PLAN, INNOCENT). The bilateral IFG and the right TPJ showed stronger activation for relevant compared to irrelevant items for those items that could be known by the respective participants.

Table 8.

Results for the whole brain analysis (p_{FWE} = .05, cluster extend threshold = 20 voxel) and the small volume corrections (12mm sphere) in prior defined ROIs (cf., Gamer, 2011).

Whole brain analysis		**x**	**y**	**z**	**Z**	***p* corrected**
Guilty action						
	left IFG	-36	24	-2	6.76	< 0.001
		-44	18	24	5.90	< 0.001
		-52	32	-4	5.86	< 0.001
	right IFG	34	26	-4	6.97	< 0.001
		42	8	34	5.07	0.017
	right SMA	4	24	52	6.51	< 0.001
	left TPJ	-62	-50	32	5.40	0.004
	right TPJ	52	-46	32	5.57	0.002
		66	-40	34	5.42	0.003
	right MFG	48	18	28	5.35	0.004
Guilty intention						
	right IFG	52	22	-6	5.26	0.007

Innocent					
left SMA	-8	16	52	6.31	< 0.001
left IFG	-48	20	10	5.95	< 0.001
right IFG	40	22	-6	5.73	0.001
ROI analysis (small volume correction)					
Guilty action					
left IFG (-44 19 1)	-36	24	-2	6.76	< 0.001
right IFG (39 23 -10)	34	26	-4	6.97	< 0.001
right TPJ (60 -48 30)	52	-46	32	5.57	< 0.001
right MFG (35 44 23)	34	44	18	3.55	0.036
Guilty intention					
left IFG (-44 19 1)	-42	24	0	5.13	< 0.001
right IFG (39 23 -10)	48	22	-4	5.23	< 0.001
right TPJ (60 -48 30)	58	-44	22	3.64	0.028
right MFG (35 44 23)	-	-	-	-	-
Innocent					
left IFG (-44 19 1)	-48	20	10	5.95	< 0.001
right IFG (39 23 -10)	40	22	-6	5.73	< 0.001
right TPJ (60 -48 30)	66	-40	34	3.93	0.011
right MFG (35 44 23)	-	-	-	-	-

Note. FWE corrected within restricted volume of interest (voxel coordinates).

Only the ACT group showed the predicted difference in activation between item types in the right MFG. The whole brain analysis additionally revealed significant activation differences for the ACT and the INNOCENT group, based on the contrast for relevant minus irrelevant items for known details (see upper part of Table 8). For the ACT group activation differences in the right supplemental motor area (SMA) and in the left TPJ were found, whereas the INNOCENT group revealed significant activations in the left SMA.

As shown in Figure 13, the extracted contrast estimates for the contrast relevant minus irrelevant details differed significantly between known and unknown details in the PLAN and INNOCENT group in most ROIs (PLAN: right IFG: $t(19) = 2.32$, $p = .031$, $d = 0.76$; left IFG: $t(19) = 3.25$, $p = .004$, $d = 0.85$, right TPJ: $t(19) = 2.39$, $p = .027$, $d = 0.73$; INNOCENT: right IFG: $t(19) = 3.27$, $p = .004$, $d = 1.07$; left IFG: $t(19) = 2.51$, $p = .021$, $d = 0.76$, right TPJ: t(19) = 2.53, $p = .021$, $d = 0.75$). In both groups no significant differences between known and unknown items were found for the right MFG. A one-way ANOVA with the between subjects factor group (ACT, PLAN, INNOCENT) was calculated for the contrast known relevant minus irrelevant details for each ROI. This analysis revealed a main effect of the intercept within most ROIs (right IFG: $F(1,57) = 42.93$, $p < .001$, $f = 0.19$; left IFG: $F(1,57) = 34.28$, $p < .001$, $f = 0.24$, right TPJ: $F(1,57) = 20.18$, $p < .001$, $f = 0.14$). For the right MFG the intercept reached only trend level, $F(1,57) = 3.98$, $p = .051$, $f = 0.22$. No other effects were significant.

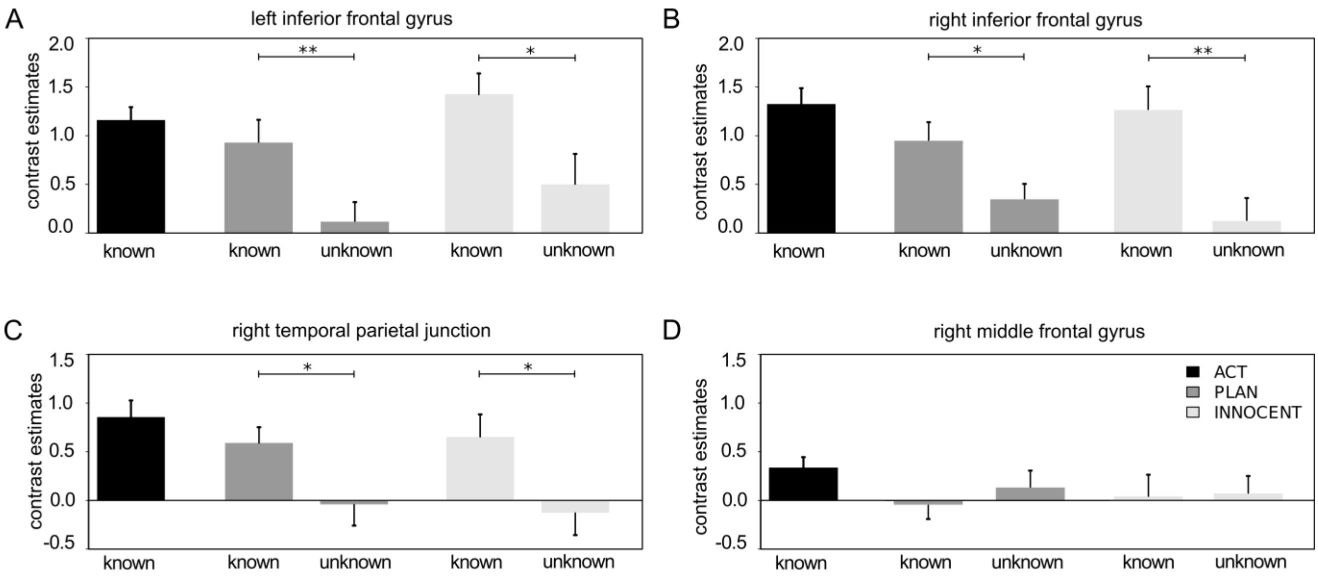


Figure 13. Mean contrast estimates (Study II) for the contrast relevant minus irrelevant details in the left IFG (A), right IFG (B), the right TPJ (C), and the right MFG (D) for known and unknown details. Error bars indicate standard error of the mean.

To examine potential differences between specific groups, the known details in the ACT group were reduced to details known by the PLAN or INNOCENT group, respectively. For the comparison between the ACT and PLAN group the left and right SMG were defined as a priori ROIs. Small volume correction for these regions revealed stronger BOLD activity for relevant details compared to irrelevant details for the ACT compared to the PLAN group in the right SMG (peak voxel: x = 40, y = -34, z = 38, $Z = 3.94$, $p_{FWE} = .025$). To compare the ACT and INNOCENT group, small volume corrections were conducted for the bilateral amygdala and the hippocampus. These analysis revealed stronger signal changes for relevant details compared to irrelevant details in the ACT compared to the INNOCENT group in the right amygdala (peak voxel: x = 28, y = -8, z = -12, Z = 3.07, $p_{FWE} = .049$) and in the right hippocampus (peak voxel: x = 30, y = -12, z = -12, $Z = 3.93$, $p_{FWE} = .012$).

3.3.5 *Validity of the CIT*

To estimate the validity of the CIT in differentiating subjects of single conditions, separate ROC curves were calculated for the right IFG, the left IFG and the right TPJ as well as for a combined ROI including these three areas. The right MFG was excluded from this analysis as prior results revealed that activity in this region did not differ between relevant and irrelevant details in the current study. Separate ROC curves were calculated for the electrodermal responses during the CIT. The area under the ROC curve can be interpreted as a validity coefficient with values around 0.5 reflecting change discrimination. Values close to one indicate valid differentiation of the respective groups. The PLAN and INNOCENT group contained two conditions regarding their knowledge of the relevant details (known/ unknown), while the ACT group had only one condition (known). As shown in Figure 14, separate areas under the ROC curve were calculated for each dependent measure (ROI activity, average of ROIs, and SCR amplitudes), contrasting all groups and conditions.

For the separate brain regions, a valid differentiation between subjects was possible when contrasting conditions where subjects had knowledge of crime related details with conditions where such knowledge was absent. As shown in Figure 14A – 14C, all comparisons between known and unknown conditions exceeded chance level independent of the encoding context. Only the coefficient for PLAN known and INNOCENT unknown in the left IFG did not reach an area under the ROC curve above chance level. Validity coefficients for similar knowledge conditions (i.e., both conditions are known or unknown) never exceeded chance level.

Similar results were found for a combined brain ROI (Figure 14D) and the electrodermal responses (Figure 14E). Again, a significant differentiation between two conditions was possible when comparing known and unknown conditions and coefficients for identical knowledge conditions did not exceed chance level.

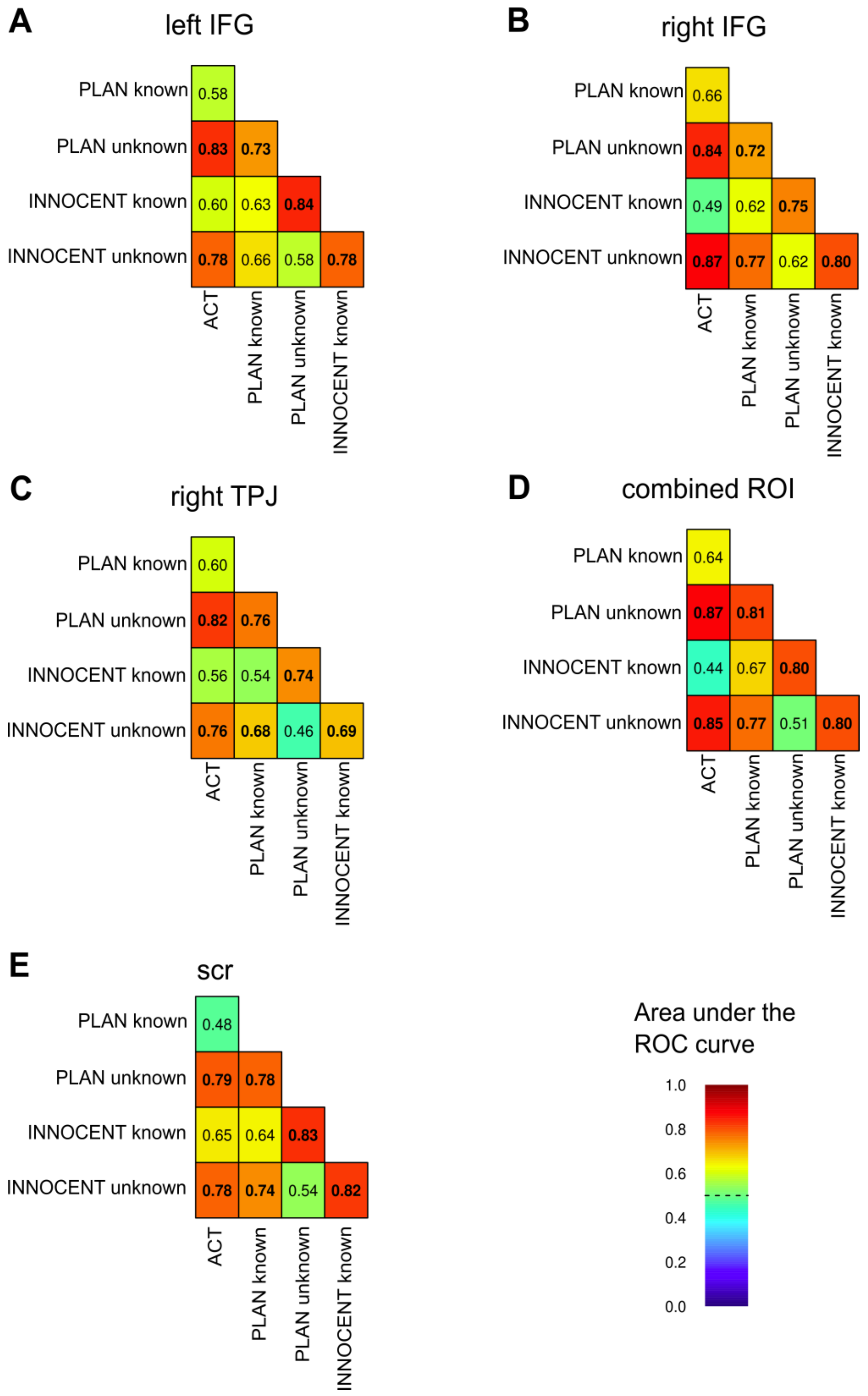


Figure 14. Areas under the ROC curve (Study II) for the left IFG (A), the right IFG (B), the right TPJ (C), a combined ROI including these brain regions (D) and electrodermal responses (E). Warm colors represent areas close to 1 and cold colors represent areas close to 0. Precise values of the areas under the ROC curve are given with bold letters representing significant differences from a chance classification of 0.5.

3.4 Discussion

The current study aimed at examining the modulation of neural activity in a CIT that asked for relevant details of a realistic mock-crime after manipulation of the subjects' knowledge and encoding context for these details. Therefore, three groups (ACT, PLAN, INNOCENT) were built to enable comparisons across conditions that differ regarding the emotional arousal during information encoding. In addition, the PLAN and INNOCNENT group remained uninformed for half of the relevant details, which enabled traditional comparisons between guilty and innocent subjects. As the CIT basically resembles a memory test for crime-related information, explicit memory is a crucial factor for a person's performance in the test. The percentage of correctly remembered relevant details directly after the CIT did not differ between groups. For the ACT group it amounted to 85.6%, for the PLAN group to 86.3%, and for the INNOCENT group to 85.0%. Thus, all groups showed comparable memory performances for details that were presented to them, independent of the respective context during information encoding. Subjective nervousness ratings confirmed that groups differed regarding their emotional arousal during their respective task. Thus, persons who committed the mock-crime reported increased nervousness compared to innocent subjects who completed an errand. The guilty group that only planned the commitment of a mock-crime reported the lowest scores, even though these subjects were not aware that they would not fulfill the mock-crime. However, when rating the subjective nervousness after the CIT, PLAN subjects already knew that they would not have to commit the mock-crime which might explain their comparably low ratings.

The three groups revealed no differences in their electrodermal response pattern after presentation of known relevant details compared to irrelevant details, which is in line with their comparable memory performance. Importantly, no differences between relevant and irrelevant details were found for the PLAN and the INNOCENT group, if restricting the analysis to unknown details. Thus, as reported in previous literature the risk of false positive results for uninformed persons seems to be small (Ben-Shakhar & Elaad, 2003). These findings further support recent studies that reported absent differences in physiological responses between guilty subjects and informed innocents (e.g., Gamer et al., 2010). Additionally, these findings replicate Meijer and colleagues (2010), who reported that the CIT can be used to detect guilty knowledge in persons who only intended to commit a crime as shown by. Thus, memory for relevant details in the second study enabled detection of this knowledge in a CIT independent of the encoding context. In contrast to these results, recent work from Elaad and colleagues (Elaad, 2013a, 2013b; Zvi, Nachson, & Elaad, 2012) reported weaker autonomic responses in informed innocents as compared to guilty subjects. The authors explain these differences by the importance of motivational aspects for performance in a CIT. However, group differences in depth of information processing in the studies by Elaad and colleagues could also account for the different physiological response pattern. As shown by Gamer and colleagues (2010) differences in strength of memory encoding between guilty subjects and informed innocents were most pronounced in a delayed CIT application. Since subjects in the second study of the current thesis showed no group differences in their delayed memory performance, similar depth of encoding can be assumed. Thus, further research is needed to clarify to what extend psychological aspects like motivation and reward influence the memory in a CIT and if there are specific differences for single physiological measures.

Regarding the fMRI data, increased activity was found after presentation of known crime relevant details in the bilateral IFG and the right TPJ independent of the respective group (ACT, PLAN; INNOCENT). These findings are broadly in line with a recent meta-analysis on the CIT (Gamer, 2011b). Interestingly, in the current study the right MFG was only found to be differentially activated by relevant and irrelevant items in the ACT group. Since previous CIT studies that reported activity changes in this brain region based on rather artificial stimuli like playing cards (cf., Gamer, 2011b), further research is necessary to specify under which circumstances neural activity changes in the right MFG occur during the CIT. Overall, the involvement of a ventral fronto-parietal network was replicated after presentation of crime relevant details during a mock-crime based CIT. Since no group differences were observed in these regions, the responses seem to occur independently of contextual aspects of information processing, like enactment or criminal versus non-criminal contexts of information encoding. This finding supports the idea that the detection of a salient detail (i.e., the crime-related information) among irrelevant stimuli is the key mechanism generating the observed neural response pattern in these brain regions. Thus, the performance during the CIT strongly depends on memory for the relevant details and deception is not necessary for detection of concealed information using this approach. Further evidence for this idea was delivered by Gamer and colleagues (2012) who reported activation of the ventral fronto-parietal network during a CIT under non-deceptive conditions (Gamer et al., 2012). However, unpublished data by Suchotzki and colleagues compared responses of guilty subjects who admitted commitment of one mock-crime and denied commitment of a second mock-crime. They found that deception enhanced activity of the neural CIT network, but that SCR responses during the CIT were less affected by deception (Suchotzki, Verschuere, Peth, Crombez, & Gamer, under review). Further research is necessary to clarify sensitivity differences between measures and their susceptibility to different psychological processes (e.g., recognition, deceptive responding) in the detection of concealed knowledge.

An exploratory whole brain analysis revealed additional areas of interest for the ACT and the INNOCENT group. Both groups showed an activation increase in the supplemental motor area (SMA) after presentation of relevant details. The SMA was repeatedly reported to be involved in orienting attention towards salient stimuli (Downar, Crawley, Mikulis, & Davis, 2001; Linden et al., 1999) and other fMRI-based CIT studies reported involvement of this brain region as well (e.g., Gamer, Bauermann et al., 2007; Gamer, Klimecki et al., 2012). Additionally, for the ACT group activation in the left TPJ was found and this region was previously reported to be involved in the detection of task relevant changes (Downar et al., 2001). Future research has to reveal if these regions are consistently involved into the recognition of salient, relevant details during a CIT.

To explore differences between groups, models restricted to the same known details were used to enable direct comparison under identical conditions. As hypothesized, the right SMG showed stronger activation for the ACT compared to the PLAN group, providing evidence for an enactment effect (Russ et al., 2003). Furthermore, as hypothesized, the right amygdala was stronger activated for the ACT compared to the INNOCENT group which is assumed to reflect differences in the emotional value of relevant details (cf., Costafreda et al., 2008) that were either encoded in an arousing mock-crime scenario (ACT group) or during a neutral task (INNOCENT group). These findings are first hints that neural activations in specific brain regions during the CIT differentiate between groups that encoded the same information in different contexts. The subjective nervousness ratings revealed that differences in emotional arousal during information encoding could be the basis for these differences in neural activity. However, the current effects are rather small and further investigations are necessary to clarify to what extent activity in these brain regions might allow for validly differentiating between groups.

The ROC analyses of the fMRI data indicate a valid differentiation between single groups when comparing different knowledge conditions (i.e., known and unknown). The combined ROI including the bilateral IFG and the right TPJ revealed an average classification accuracy of 0.82. In contrast, the electrodermal responses reached only an average classification accuracy of 0.79 in the current study. Overall, the areas under the ROC curve in the current study were comparable to coefficients reported for autonomic measures (cf., Ben-Shakhar and Elaad, 2003). Thus, from an applied perspective, the fMRI-based CIT did not outperform the traditional CIT and the higher amount of investment for measurement of neuroimaging data did not result in improved accuracy rates.

A lack of research exists regarding longer time intervals between encoding of relevant details and the CIT application. Hira (2003) reported temporal stability of a P300-based CIT when the test was delayed up to one year. However, this study used only one highly salient detail (the stolen good). Moreover, it is currently unknown how well different aspects of a crime are remembered over such long time periods. In the current study, the memory test for relevant crime details was repeated by an internet questionnaire six months after study participation. Dropout rate was acceptable (30.0%) and all groups showed comparable memory performance after this comparably long time period. The percentage of correctly remembered relevant details amounted to 75.0% for the ACT, 72.2% for the PLAN, and 76.1% for the INNOCENT group. Implications of this rather stable memory are somewhat limited because no CIT was applied after six month, however these results deliver evidence that memory for relevant details shows some degree of persistence over time. To further investigate this topic, future research should apply more CIT interrogations after longer time intervals.

To conclude, this study examined for the first time the influence of encoding context on neural and autonomic activity in a CIT examination using a realistic mock-crime procedure. Taken together, the current findings demonstrated that persons who got knowledge about crime relevant details show activations in a ventral fronto-parietal network, that seems to detect salient information among irrelevant details. This reaction occurred independent of differences in the encoding context (criminal versus non-criminal) and supports the idea that memory for relevant details is the crucial process for performance during the CIT instead of deception. Thus, there is a high risk for informed innocents to be falsely classified as guilty. Nevertheless, the current results also indicate that it seems to be possible to detect concealed information based on intentions before they are put into action.

4. General discussion

The NRC (2003) raised concerns about the influence of emotional and motivational factors on the detection of deception, which mainly apply to approaches like the CQT. So far, the influence of emotional factors on memory for relevant details of a crime during the CIT has rarely been investigated. Besides its basis on scientific research and theory (cp., Verschuere & Ben-Shakhar, 2011), the external validity of the CIT was repeatedly debated and concerns about memory for relevant details in real-life applications were raised (Honts, 2004). Basic research supports the idea that emotional arousal modulates memory for details of a scenario (e.g., Adolphs et al., 2001; Christianson, 1992). In addition, real-life settings are assumed to induce higher emotional arousal in a person, but implications of these findings for CIT settings are unclear.

The current thesis investigated the emotional modulation of information encoding during a mock-crime and retrieval of relevant crime details during the CIT by answering the following research questions. First, does emotional arousal during the mock-crime influence autonomic responses during the CIT? Second, are neural activations during the CIT modulated by the encoding context? In addition, the current thesis investigated whether the application of ocular measures enables the detection of concealed knowledge during the CIT. The following sections link the respective study findings reported in the current thesis with the open research questions, based on the dependent measures used during the CIT. In addition, limitations of the current studies and potential implications for the CIT in the field are discussed. Finally, an outlook on future research questions is presented.

4.1 Influence of emotional factors on the CIT

To answer the previously defined research questions, two studies were conducted that aimed at investigating the emotional modulation of memory for relevant details of a mock-crime under conditions that better resembled real-life situations. For example, the relevant details of the mock-crime were only incidentally encoded during the preparation or the enactment of the mock-crime and not explicitly learned prior to the CIT, as done by previous laboratory studies (cf., Ben-Shakhar and Elaad, 2003). To further approximate field conditions, the first study included manipulations of the amount of emotional arousal during commitment of the mock-crime and the length of delay between mock-crime and CIT application. Therefore, additional emotional arousal was induced in half of the guilty subjects and the CIT measurement occurred either immediately after the mock-crime or two weeks later. Subjective and objective nervousness measures revealed that the arousal manipulation was successful and subjects in the arousal induction group were more nervous compared to the group without arousal induction.

The second study manipulated the encoding context by asking subjects to plan and/or fulfill either a mock-crime or an errand. This enabled comparisons of neural activity during the CIT between guilty subjects and informed innocents or subjects who only intended to commit the mock-crime, respectively. Subjective nervousness ratings after the CIT revealed stronger emotional involvement for the subjects who committed the mock-crime compared to the other groups.

4.1.1 Emotional modulation of explicit memory for relevant details

Memory for the relevant details of a crime is a necessary condition for detection of concealed information during the CIT. The first study included central details that were directly involved in the mock-crime (i.e., a money box subjects had to open), as well as peripheral details that were present in the scenario without an active involvement (i.e., a picture on the wall). Regarding the explicit memory for the relevant details, an advantage in memory for central compared to peripheral details was found and in general, memory for relevant details was diminished after two weeks. Interestingly, the decreased memory for peripheral details was more pronounced in the arousal induction group. These results support findings from basic research that emotional arousal primarily reduces memory for peripheral information (Adolphs et al., 2001). However, in contrast to previous findings in basic research (cf., Kensinger, 2009; Steblay, 1992) no enhanced recognition of central details in the arousal induction group occurred, probably due to a ceiling effect.

In the second study, no differences in explicit memory were found between the groups (ACT, PLAN, INNOCENT). Thus, independent of the encoding context (criminal versus non-criminal; intention versus enactment) subjects remembered on average between 85.0% and 86.3% of the relevant details. A delayed memory test after six month was conducted in 70% of the original study sample and again revealed no differences between groups. In general, a decrease in memory occurred over time and resulted in average recognition rates between 72.0% and 76.2%. In line with the explicit memory test immediately after the CIT no influence of encoding context was found.

In sum, no differences in memory for relevant details was found between guilty subjects and informed innocents or guilty subjects who only planed the mock-crime. Thus, the explicit memory for relevant details was not modulated by the encoding context in the second study. Nevertheless, findings from the first study revealed that emotional arousal during the mock-crime diminished memory for peripheral details, while central details were remembered well. Thus, emotional arousal during encoding of relevant details modulated the recognition of specific details of the mock-crime and not the general memory performance.

4.1.2 Emotional modulation of autonomic responses during the CIT

For the first study, the electrodermal, respiratory, finger pulse and heart rate responses during the CIT were measured. In line with the explicit memory performance, the physiological responses during the CIT were stronger for central compared to peripheral details. Moreover, a combined score of electrodermal, respiratory and heart rate responses enabled a valid differentiation between guilty and innocent persons two weeks after the mock-crime when the analysis was restricted to central crime details. Additionally, increased emotional arousal during information encoding strengthened the physiological responses during the CIT over time. In contrast, chance level classification was reported in the delayed CIT if the analysis was restricted to peripheral details.

The diminished standard difference scores for heart rate, respiration and finger pulse responses at the delayed CIT measurement occasion in the first study are consistent with the reduced explicit recognition rate after two weeks. This finding replicated the temporal decline for different physiological measures reported by Gamer and colleagues (2010). However, in contrast to the latter study, no significant differences between electrodermal responses at the immediate and the delayed measurement occasion were found in the first study. Other realistic mock-crime studies reported heterogeneous results regarding differences in skin conductance responses of guilty subjects over time (Carmel et al., 2003; Nahari & Ben-Shakhar, 2011). Therefore, further research is necessary to clarify under which circumstances electrodermal responses are affected by a delayed CIT application and if this measure might be less influenced by situational conditions as compared to respiratory and cardiovascular measures.

Overall, emotional arousal in the first study was found to increase the detection of concealed central details during the CIT. In addition, the data from the first study revealed differences in sensitivity for increased emotional arousal during encoding of the relevant details in single autonomic measures. Thus, respiratory responses and heart rate reactions showed trends for an interaction effect between time of test and arousal induction, with no decrease in the standard difference scores for these measures in the arousal induction group after two weeks. These findings are in line with the existing CIT literature that reported respiratory measures to be especially sensitive for emotionally arousing crime aspects (Suzuki et al., 2004) and to enable a better detection of concealed information under field conditions (Elaad et al., 1992). However, since the current findings reached only trend level, further research is needed to investigate possible differences between autonomic measures in their sensitivity for emotional crime aspects.

The second study measured only electrodermal responses during the fMRI-based CIT. In line with the first study, known relevant details resulted in an increased electrodermal response compared to irrelevant details. No such effect was found in uninformed persons (i.e., responses of PLAN and INNOCENT groups for unknown relevant details). The encoding context seemed to be of little influence, as differences between groups were only found at trend level. Therefore, even persons who got knowledge about relevant details in a non-criminal context showed response patterns comparable to guilty subjects. This finding is in line with previous research that failed to find differences in autonomic responses between guilty subjects and informed innocents (e.g., Gamer et al, 2010). In addition, the PLAN group did not differ in their electrodermal response to known relevant details compared to the guilty ACT group. This finding is in line with previous research, reporting valid detection of intentions to commit a mock-crime based on electrodermal responses (Meijer et al., 2010) or reaction times (Noordraven & Verschuere, 2013). Important for field CIT applications, legal consequences might become more relevant when considering the usage of the CIT to detect intentions. Usually a person's innocence is assumed unless a crime was conducted and the person's guilt is proven. Therefore, ethical concerns and legal limitations have to be discussed before potential field applications are considered, for example regarding the prevention of terrorist attacks.

The validity coefficients for electrodermal responses in the second study enabled a valid differentiation between informed and uninformed persons, independent of the context of information encoding. This finding further supports the idea that memory for relevant details is a sufficient condition for valid CIT results. However, the first study showed for the first time that emotional arousal has a positive effect on the detection of concealed information during a CIT with autonomic measures. In addition, no decrease in validity of the CIT was found for a delayed measurement after the induction of emotional arousal during the mock-crime. Thus, memory for the relevant details is assumed to be the main mechanism to detect concealed information during the CIT. This effect seems to be rather insensitive for specific conditions of information encoding in a person (i.e., the encoding context). Nevertheless, emotional arousal was found to influence responses during the CIT over time and differences in sensitivity for emotional modulation of recognition across autonomic measures were revealed. Thus, emotional aspects are assumed to modulate the basic memory mechanisms underlying the CIT.

4.1.3 Emotional modulation of neural activity during the CIT

The second study of the current thesis investigated the influence of encoding context on a person's memory during an fMRI-based CIT. In line with prior CIT research using neuroimaging (cf., Gamer, 2011b), the act of information concealment modulated activity in a ventral fronto-parietal network consisting of the bilateral IFG and the right TPJ. Comparable to the explicit memory performance, groups (ACT; PLAN; INNOCENT) did not differ with respect to neural activity in these brain regions for known relevant compared to irrelevant details. These findings support the idea that the general recognition of relevant details among irrelevant details during the CIT is critical for a person's performance during the test, independent of the encoding context. However, the identified ROIs are not exclusively involved in memory detection processes. For example, the IFG has been linked to detection of relevant details in a train of irrelevant ones (Kiehl et al., 2001) as well as to improved retrieval of relevant details from memory (Iidaka et al., 2006). Furthermore, the TPJ was reported to be activated by infrequent changes in the environment (Downar et al., 2000, 2002). In addition, activity in IFG and TPJ were found to reflect detection of unexpected or rare events independent of sensory modality or response demands (Corbetta & Shulman, 2002). Thus, the regions activated during the concealment of information in the CIT reflect rather general functions of the brain instead of deception specific functions. However, deception was recently reported to enhance activity in this network (Suchotzki et al., under review). In the current study, all subjects tried to hide their knowledge. Even subjects in the

INNOCENT group were informed that they could get a monetary reward after successful classification as an innocent person. However, they realized during the CIT that some of the presented answer options were known by them from the errand and they had to deny that knowledge. Further research is necessary to further explore the influence of deception on the detection of concealed information during the CIT.

The increased activation in the bilateral IFG and the right TPJ for known relevant compared to irrelevant details in the second study were accompanied by increased activation of the SMA after presentation of known relevant details in the ACT and the INNOCENT group. This brain region was reported in other CIT studies as well. For example, Gamer and colleagues (2007) found that response times and SCR amplitudes during the CIT positively correlated with activation in the SMA and the right IFG (Gamer et al., 2007). The couplings between these brain regions and behavioral or autonomic measures were assumed to underline the modulation of sympathetic arousal in the CIT. Among other functions, the SMA also reflects the orienting of attention towards salient stimuli (Linden et al. 1999; Downar et al., 2001). In addition, the ACT group showed increased activation in the left TPJ after presentation of relevant details. This region was previously reported to be involved in the detection of task relevant changes (Downar et al., 2001). As only the ACT group showed increased activation in this region, this finding might be a hint for an increased response to the relevant details in this group compared to the other groups. In addition, the ACT group was the only one that showed increased activity in the right MFG in the second study. However, future research will have to investigate under which conditions specific brain regions are involved in the recognition of relevant details during the CIT and whether these regions should be taken into account as ROIs during fMRI-based CIT investigations.

Based on the different encoding contexts, group differences regarding specific ROIs were hypothesized. In detail, the ACT and INNOCENT group were assumed to differ regarding their emotional arousal during encoding of the relevant details, because subjects of the former group committed a mock-crime whereas participants of the latter condition fulfilled an errand. Subjective nervousness ratings revealed that the guilty ACT group task was more aroused compared to the INNOCENT group task. Based on research on the emotional modulation of memory (cf., Costafreda et al., 2008), increased neural activation in amygdala and hippocampus after presentation of the relevant details were hypothesized for the ACT compared to the INNOCENT group. To enable a proper comparison, an analysis restricted to details known in both groups was conducted. Here, increased activity in the right amygdala and the right hippocampus were found for the ACT compared to the INNOCENT group when contrasting known relevant and irrelevant details. Nevertheless, no differences between groups were found regarding neural activation in the CIT network. Therefore, informed innocent subjects could not be differentiated from guilty subjects based on these ROIs and have a high risk to be falsely classified as guilty when relying on these previously specified brain regions (Gamer, 2011b).

The PLAN group was expected to encode the relevant details less deep compared to the ACT group, because these subjects only read about the relevant details instead of committing the real mock-crime. This hypothesis was based on basic research that reported an advantage in memory for information processed during an action in contrast to passive reading – the so called enactment effect (Russ, 2003). In line with these results, neural activity differences in the SMG were hypothesized between the ACT and the PLAN group. Again, an analysis restricted to the relevant details known in both groups was conducted to enable this comparison under similar conditions. As hypothesized, the ACT group showed increased activity in the right SMG compared to the PLAN group when contrasting relevant and irrelevant details. This finding supports the assumption that information encoded during the real commitment of a mock-crime is processed differently compared to information based on the planning of the mock-crime. Indeed, basic research reported the SMG to be involved in action execution, simulation and observation (Grèzes & Decety, 2001) as well as .

Importantly, the PLAN and ACT group did not differ regarding their neural activity in the CIT network. Thus, subjects who intended to commit a mock-crime can be detected with an fMRI-based CIT.

In sum, activation of the CIT network reveals whether a person has specific knowledge about relevant details, but this response occurs independent of the encoding context and leads to a high risk for informed innocent persons. Future research is necessary to replicate the current findings and to investigate in more detail whether activation in brain regions like the amygdala, the hippocampus and the SMG might be used as additional predictor to identify under which conditions a person got knowledge about specific details of a crime. So far, autonomic measures were not able to cover such specific needs.

Regarding the detection accuracy, the areas under the ROC curve for the fMRI data reached scores between 0.76 and 0.85 for the differentiation between informed and uninformed persons. The electrodermal responses in the current study reached areas between 0.74 and 0.83. Thus, autonomic and central nervous measures showed comparable validity coefficients. Incremental validity calculations could help to clarify whether these measures reflect different aspects of the same phenomenon, as already shown for other measures. For example, Gamer and Berti (2010) reported in an EKP-based CIT study that a combined score of the N200 component and electrodermal responses increased the detection of concealed information (Gamer & Berti, 2010). It is important to keep in mind that the optimal conditions for data collection might differ between autonomic measures and fMRI. Since the activity of the central nervous system during the CIT was primarily investigated with event-related brain potentials, these designs usually included a high amount of repeated stimulus presentations with short inter stimulus intervals (i.e., substantially below 5 s). For fMRI studies similar CIT designs were used. In contrast, autonomic measures are usually recorded in a setting where each stimulus is only shown once or twice with very long inter stimulus intervals (i.e., around 20 s). Thus, a possible influence of habituation effects after repeated presentation of the same stimuli during a CIT should be investigated in more detail or validity coefficients should be compared between CIT applications that are conducted under optimal conditions for the respective dependent measure (Verschuere et al., 2010).

In sum, the second study used for the first time an fMRI-based CIT in a mock-crime study. Importantly, the ventral fronto-parietal CIT network was found to respond to concealed information, but it was not sensitive to the encoding context.

4.1.4 Application of eye-movements and eye-blinks during the CIT

In addition to the traditionally applied autonomic measures, data on ocular measures were recorded during the first study to covertly observe behavioral responses. For the first time, it was shown that the number of fixations and the average fixation duration enabled a valid differentiation between guilty and innocent persons in a standard CIT setting. This effect was most pronounced for central details, during stimulus presentation and 5 s after stimulus-offset. In line with research on the eye-movement memory effect (Hannula et al., 2010), less exploration of previously seen stimuli (i.e., relevant details of the mock-crime) occurred during the CIT. In contrast, eye-blinks enabled a valid detection of concealed information only in the 5 s time period after stimulus-offset. These findings further support the idea that different processes explain the physiological responses during the CIT. As existing literature on the theoretical background of the CIT suggests (cf., Ben-Shakhar and Elaad, 2003), an OR could explain the changes in fixation behavior during stimulus presentation very well. The eye-blink changes were most effective in a 5 s time interval after stimulus-offset. These delayed responses could be explained by response inhibition, as suggested by CIT research using startle responses (Verschuere, Crombez, Koster, Van Bockstaele, et al., 2007). Verschuere and colleagues (2007) failed to find increased startle responses during a CIT as predicted by OR theory. Instead they found a reduced startle modulation, which further supports the contribution of inhibition to physiological responses during the CIT. Additional support for this idea was delivered by fMRI research that associated concealed information detection with activity in the bilateral IFG (cf., Gamer, 2011b), a brain region that was linked to response inhibition in previous studies as well (Aron, Robbins, & Poldrack, 2004). Taken together, the findings on ocular measures in the current study provide some evidence that concealment of information during the CIT involves different processes. A bottom-up driven OR is assumed to occur early during information processing and to enable the detection of salient information among irrelevant details during the CIT. After this primary reaction, processes that are stronger controlled by the person, like response inhibition, are assumed to be stronger involved.

One limiting factor of the current data is the complete failure to reveal crime-related memory for peripheral details using ocular measures. This is surprising, as prior research reported that peripheral information could be validly detected based on autonomic responses, at least when applying the CIT immediately after the mock-crime (Gamer et al., 2010; Nahari & Ben-Shakhar, 2010). A lack of sensitivity of the ocular measures could explain this effect. As revealed by the calculated validity estimates, gaze measures showed weaker effect sizes compared to traditionally used autonomic measures. The differential effects for central and peripheral details also mirror results from research on eyewitness memory. For example, an emotional enhancement of memory is frequently only observed for central details (cf., Christianson, 1992). These potential limitations of ocular measures as indicators of concealed information should be investigated in further detail.

Nevertheless, ocular measures can be recorded covertly without application of any equipment to a subject's body, thus the current findings are highly relevant for field applications. So far, covert respiration was the only unobtrusive measurement reported to successfully enable detection of concealed information during the CIT (Elaad & Ben-Shakahr, 2008; 2009). Other approaches, like voice stress analysis (e.g., Gamer et al., 2006) were critically discussed and their validity is strongly questioned (cf., Elaad, 2011b). In sum, ocular measures seem to be a promising approach, but the current findings require replication in future studies.

4.1.5 Limitations

Overall, the results of the current thesis are limited to some extent. Similar to other laboratory studies in this domain, transfer of the current findings to real-life applications is relatively unclear. Although we tried to construct more field-like conditions using incidental encoding procedures, manipulations of emotional arousal and delayed CIT applications, all subjects gave written informed consent to participate and they all knew that a CIT would be part of the study at some point. In addition, they received money for participation and followed standardized instructions. Thus, these persons did not decide by themselves to commit a crime and it remains unclear whether persons who planned and executed real crimes by themselves would respond differently in a CIT. Additionally, subjects in the current studies were young and healthy adults, most of them college students of different fields. It is questionable whether persons who commit crimes in real-life are comparable in their memory for relevant details and show similar physiological responses compared to the investigated subjects.

Verschuere and colleagues (2007) investigated prisoners with an autobiographic CIT and reported generally reduced autonomic reactivity in this population compared to a control group. Interestingly however, concealed information was still validly detected in both groups and showed no significant difference between prisoners and controls (Verschuere, Crombez, Koster, & De Clercq, 2007). This finding supports the idea that the CIT would work in real-life settings as well, but due to the lack of field studies the external validity of the current findings is still limited.

Following this point, the emotional arousal in the first study was artificially induced and the external validity of this manipulation is unclear. However, physiological and subjective arousal measures verified that the mock-crime was stressful for all groups of guilty examinees. Thus, the additional arousal manipulation was successful although the effect of the arousal induction on heart rate responses and subjective nervousness ratings was comparably small. Interestingly, the arousal induction had a longer lasting effect and resulted in a relative heart rate increase that was more pronounced after participants completed the mock-crime and returned to the examination room. The relatively small effect of the arousal induction on heart rate responses might be related to a ceiling effect given that all guilty examinees showed a strong heart rate increase during the mock-crime as compared to baseline (more than 25 bpm on average). Potentially, this effect was also modulated by the experimental setting that could not be perfectly controlled with respect to other arousal inducing factors. For example, employees of the institute might have passed the room when the mock-crime took place or they chatted in the corridor in front of the storeroom, which might have enhanced stress in all examinees. Resulting questions for future research are for example to what extent other factors in the environment can influence the participants arousal and whether the pattern of arousal that was induced in this study is comparable to the arousal experienced by offenders who routinely commit crimes (Hira & Furumitsu, 2009; Verschuere, Crombez, Koster, & De Clercq, 2007).

In addition, the time delay between commitment of the mock-crime and CIT application is an important point. Results from studies that used a delay of one to two weeks (e.g., Carmel et al., 2003; Elaad, 1997; Gamer et al., 2010; Nahari & Ben-Shakhar, 2011) might not generalize to realistic CIT applications that can occur several months or even years after the crime. So far, two studies reported a temporal stability of the P300-based CIT when the test was delayed up to one month (Hu et al., 2012) or one year (Hira, 2003), but these studies only used one highly salient CIT item (the stolen good). Thus, it is currently unknown whether other aspects of a crime could also be reliably detected in a CIT examination after such a long period of time. During the second study of the current thesis, an explicit memory test for all relevant details (i.e., 24 questions) was conducted after six month with an internet questionnaire. These data revealed decreased memory over time, but the relevant details were still remembered clearly above chance level (74.4% average recognition rate). However, because no repeated collection of physiological responses occurred for these subjects, changes in physiological responding are unclear. Interestingly, studies on implicit memory reported preserved priming effects for picture recognition over 48 weeks (Cave, 1997) and for picture fragment identifications even over 17 years (Mitchell, 2006). In addition, basic research on memory reported a better memory for emotional pictures compared to neutral pictures one year after encoding, accompanied by increased activation in the amygdala and hippocampus during retrieval (Dolcos, LaBar, & Cabeza, 2005). These findings indicate that the CIT might be applicable even after longer time periods but future studies need to examine whether this is really the case.

4.2 Implications for the CIT usage in the field

Overall, memory for the relevant details is the key mechanism that explains a person's performance during the CIT. The second study of the current thesis revealed that memory for relevant details is detected in a CIT independent of the encoding context. Thus the CIT is insensitive for the leakage of information and persons who got knowledge about the relevant details in a non-criminal context have a high risk to be classified as guilty. Interviews prior and after the CIT should be done to assure that suspects can explain their knowledge (cf., Osugi, 2011). It is very important that no details are presented to a suspect, for example during previous interviews or in newspapers.

The first study of this thesis showed that emotional arousal strengthened physiological responses during the CIT and improved the detection of concealed knowledge of central crime details over time. Thus, persons who are strongly aroused during the commitment of a crime will better remember central aspects of a crime and easier forget about peripheral details. Accordingly, practical investigators should choose details for the CIT that were directly involved in the crime and actively handled by the suspect (e.g., the stolen good). These details are better remembered compared to peripheral details of the crime scene (e.g., a picture on the wall). If there is a lack of adequate details for a specific crime, it is not advisable to construct a CIT based on details that were not directly involved in the crime.

Differences in sensitivity for increased emotional arousal during the mock-crime were found at trend level for the heart rate and respiratory responses in the first study and the corresponding combined physiological score reflected a stronger response over time after emotional arousal induction. Thus, real-life investigations should take into account not only electrodermal responses to increase CIT validity.

In addition, eye-movements and eye-blinks enabled a valid detection of concealed information in the first study. This finding could be useful in field applications that require a covert response recording. However, the detection was limited to central crime details and validity scores were smaller compared to traditionally used autonomic measures. In contrast, the detection accuracy of the fMRI-based CIT in the second study reached areas under the ROC curve that were similar to autonomic measures. These findings indicate that the use of neuroimaging techniques is not advisable for field investigations since the measurement and data analysis are rather expensive and more time-consuming compared to autonomic measures.

4.3 Outlook

A general problem for the application of all deception detection techniques is their susceptibility to countermeasures. Especially guilty subjects may often attempt to systematically manipulate their response pattern during the test (for a review see Ben-Shakhar, 2011). Frequently applied countermeasures are either physical (e.g., movements of toes or fingers) or mental manipulations (e.g., remembering an emotional arousing event) that could result in a faked response increase for irrelevant details during the CIT. For example, physical countermeasures were shown to decrease the detection of concealed information in an autobiographic, fMRI-based CIT (Ganis et al., 2010). However, the physical movements were detected based on increased activity in the primary motor cortex. Heterogeneous results are reported for studies using autonomic measures during the CIT, depending on the type of countermeasure used and the autonomic measure recorded (e.g., Ben-Shakhar & Dolev, 1996; Elaad & Ben-Shakhar, 1991). For example, electrodermal responses were more sensitive for the usage of physical countermeasures compared to respiratory measures, but both measures were more resistant against mental countermeasures (e.g., Elaad and Ben-Shakhar, 2009). ERP studies reported a new component, the P900, to be an indicator for countermeasure usage during the CIT (Rosenfeld & Labkovsky, 2010). However, recent findings showed that the P900 effect was limited to very specific conditions and only occurred if two or three details out of five details were manipulated during the test (Labkovsky & Rosenfeld, 2012). Overall, these results might be limited regarding their transfer to real-life applications, because most of the CTP studies used autobiographical information (e.g., date of birth).

To further investigate the influence of countermeasures on the CIT is an important challenge for future CIT research. For example, mental countermeasures should be used in an fMRI-based CIT to estimate whether their influence on a person's classification is comparable to autonomic measures. These studies should focus on mock-crime scenarios instead of autobiographic CIT designs and more realistic conditions like a delayed CIT application should be used. To investigate the validity of the CIT under conditions that optimize real-life conditions and increase the motivation of examinees to pass the test is the next step to estimate the CIT's applicability in forensic cases.

4.4 Conclusion

In sum, the current data support the CIT as a valid technique to detect concealed information. Increased emotional arousal during the mock-crime was found to strengthen autonomic responses for relevant crime details during the CIT. This effect was especially pronounced for central details and remained stable over a time delay of two weeks. These data indicate that a higher external validity in the laboratory does not necessarily lead to decreased validity coefficients of the CIT as has been hypothesized previously (Honts, 2004). In addition, a previously defined ventral fronto-parietal brain network (cf., Gamer, 2011b) enabled a valid detection of concealed information independent of the encoding context.

In sum, the current data support the assumption that recognition of relevant crime details is the key mechanism that drives the response pattern in the CIT on the autonomic and neural level. Nevertheless, specific emotional factors (i.e., increased emotional arousal during encoding) were found to modulate the autonomic responses during the test in guilty subjects. Thus, emotional arousal was not necessary for detection of concealed information in the CIT, but can increase the detection accuracy under specific conditions. Further research is necessary to better understand how and when emotional factors influence the validity of the CIT and what other factors are important for the encoding and recognition of crime-related information.

5. References

Adolphs, R., Denburg, N. L., & Tranel, D. (2001). The amygdala's role in long-term declarative memory for gist and detail. *Behavioral Neuroscience*, *115*(5), 983–992. doi:10.1037//0735-7044.115.5.983

Allen, J. J., Iacono, W. G., & Danielson, K. D. (1992). The Identification of Concealed Memories Using the Event-Related Potential and Implicit Behavioral Measures: A Methodology for Prediction in the Face of Individual Differences. *Psychophysiology*, *29*(5), 504–522. doi:10.1111/j.1469-8986.1992.tb02024.x

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: a reprocessing effect in face perception. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*(4), 997–1010.

Ambach, W., Stark, R., Peper, M., & Vaitl, D. (2008). An interfering Go/No-go task does not affect accuracy in a Concealed Information Test. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *68*(1), 6–16. doi:10.1016/j.ijpsycho.2007.11.004

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*(4), 170–177. doi:10.1016/j.tics.2004.02.010

Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of Mathematical Psychology*, *12*(4), 387–415. doi:10.1016/0022-2496(75)90001-2

Barry, R. J. (1996). Preliminary process theory: Towards an integrated account of the psychophysiology of cognitive processes. *Acta Neurobiologiae Experimentalis*, *56*(1), 469–484.

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

Ben-Shakhar, G., & Dolev, K. (1996). Psychophysiological detection through the guilty knowledge technique: Effect of mental countermeasures. *Journal of Applied Psychology, 81*(3), 273–281. 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*(5), 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. *The Journal of Applied Psychology*, *88*(1), 131–151. doi:10.1037/0021-9010.88.1.131

Ben-Shakhar, G., Gronau, N., & Elaad, E. (1999). Leakage of relevant information to innocent examinees in the GKT: An attempt to reduce false-positive outcomes by introducing target stimuli. *Journal of Applied Psychology*, *84*(5), 651–660. doi:10.1037/0021-9010.84.5.651

Ben-Shakhar, G, & Furedy, J. J. (1990). *Theories and applications in the Detection of Deception: A Psychophysiological and International Perspective*. Springer. New York.

Berntson, G. G., & Cacioppo, J. T. (2000). From homeostasis to allodynamic regulation. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.). *Handbook of Psychophysiology* (pp. 459-481). Cambridge University Press.

Berntson, G.G., Quigley, K.S., & Lozano, D. (2007). Cardiovascular Psychophysiology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson, (Eds.). *Handbook of Psychophysiology* (pp. 182-210). Cambridge University Press.

Bradley, M. T., & Ainsworth, D. (1984). Alcohol and the Psychophysiological Detection of Deception. *Psychophysiology*, *21*(1), 63–71. doi:10.1111/j.1469-8986.1984.tb02319.x

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

Bradley, M. T., MacLaren, V. V., & Carle, S. B. (1996). Deception and nondeception in guilty knowledge and guilty actions polygraph tests, *81*, 153–160.

Bradley, M. T., & Rettinger, J. (1992). Awareness of crime-relevant information and the guilty knowledge test. *Journal of Applied Psychology*, *77*, 55–59.

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

Bressler, S. L., & Ding, M. (2006). Event-Related Potentials. In M Akay (Ed.) *Wiley Encyclopedia of Biomedical Engineering*. Wiley. Hoboken NJ.

Buchanan, T. W. (2007). Retrieval of emotional memories. *Psychological Bulletin*, *133*(5), 761–779. doi:10.1037/0033-2909.133.5.761

Cahill, L., Haier, R. J., Fallon, J., Alkire, M. T., Tang, C., Keator, D., McGaugh, J. L. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(15), 8016–8021.

Carmel, D., Dayan, E., Naveh, A., Raveh, O., & Ben-Shakhar, G. (2003). Estimating the validity of the guilty knowledge test from simulated experiments: the external validity of mock-crime studies. *Journal of Experimental Psychology. Applied*, *9*(4), 261–269. doi:10.1037/1076-898X.9.4.261

Cave, C. B. (1997). Very Long-Lasting Priming in Picture Naming. *Psychological Science*, *8*(4), 322–325. doi:10.1111/j.1467-9280.1997.tb00446.x

Christianson, S.-Å. (1992). *The Handbook of emotion and memory: research and theory*. Routledge.

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

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–215. doi:10.1038/nrn755

Costafreda, S. G., Brammer, M. J., David, A. S., & Fu, C. H. Y. (2008). Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews*, *58*(1), 57–70. doi:10.1016/j.brainresrev.2007.10.012

De Winstanley, P. A., & Bjork, E. L. (1997). Processing instructions and the generation effect: a test of the multifactor transfer-appropriate processing theory. *Memory*, *5*(3), 401–21. doi:10.1080/741941392

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating?. *Behavioral and Brain Sciences*, *11*(03), 357–374. doi:10.1017/S0140525X00058027

Dolcos, F., Denkova, E., & Dolcos, S. (2012). Neural Correlates of Emotional Memories: A Review of Evidence from Brain Imaging Studies. Special Issue on "Recent Advances of Functional Neuroimaging Studies on Episodic Memories", *Psychologia, 55*, 80-111.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron, 42*(5), 855–863.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), 2626–2631. doi:10.1073/pnas.0409848102

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, *3*(3), 277–283. doi:10.1038/72991

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2001). The Effect of Task Relevance on the Cortical Response to Changes in Visual and Auditory Stimuli: An Event-Related fMRI Study. *NeuroImage*, *14*(6), 1256–1267. doi:10.1006/nimg.2001.0946

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2002). A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. *Journal of Neurophysiology*, *87*(1), 615–620.

Elaad, E. (1990). Detection of guilty knowledge in real-life criminal investigations. *Journal of Applied Psychology*, *75*(5), 521–529. doi:10.1037/0021-9010.75.5.521

Elaad, E. (2009). Effects of context and state of guilt on the detection of concealed crime information. *International Journal of Psychophysiology*, *71*(3), 225–234. doi:10.1016/j.ijpsycho.2008.10.001

Elaad, E. (2011a). Validity of the Concealed Information Test in realistic contexts. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 171–186). Cambridge: Cambridge University Press.

Elaad, E. (2011b). New and old covert measures in the Concealed Information Test. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 114–127). Cambridge: Cambridge University Press.

Elaad, E. (2013a). Differences in the readiness of guilty and informed innocent examinees to cooperate on the Guilty Action Test. *Psychophysiology*, in press. doi:10.1111/psyp.12146

Elaad, E. (2013b). Effects of goal- and task-oriented motivation in the guilty action test. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *88*(1), 82–90. doi:10.1016/j.ijpsycho.2013.02.004

Elaad, E., & Ben-Shakhar, G. (1991). Effects of mental countermeasures on psychophysiological detection in the guilty knowledge test. *International Journal of Psychophysiology*, *11*(2), 99–108. doi:10.1016/0167-8760(91)90001-E

Elaad, E., & Ben-Shakhar, G. (2006). Finger pulse waveform length in the detection of concealed information. *International Journal of Psychophysiology*, *61*(2), 226–234. doi:10.1016/j.ijpsycho.2005.10.005

Elaad, E., & Ben-Shakhar, G. (2008). Covert respiration measures for the detection of concealed information. *Biological Psychology*, *77*(3), 284–291. doi:10.1016/j.biopsycho.2007.11.001

Elaad, E., & Ben-Shakhar, G. (2009). Countering countermeasures in the concealed information test using covert respiration measures. *Applied Psychophysiology and Biofeedback*, *34*(3), 197–208. doi:10.1007/s10484-009-9090-5

Elaad, E., Ginton, A., & Jungman, N. (1992). Detection measures in real-life criminal guilty knowledge tests. *Journal of Applied Psychology*, *77*(5), 757–767. doi:10.1037/0021-9010.77.5.757

Engelkamp, J. (1998). *Memory for actions.* Psychology Press, Hove, UK.

Farwell, L., & Donchin, E. (1991). The Truth Will Out - Interrogative Polygraphy (lie Detection) with Event-Related Brain Potentials. *Psychophysiology*, *28*(5), 531–547. doi:10.1111/j.1469-8986.1991.tb01990.x

Fukuda, K. (2001). Eye blinks: new indices for the detection of deception. *International Journal of Psychophysiology*, *40*(3), 239–245. doi:10.1016/S0167-8760(00)00192-6

Gamer, M. (2010). Does the Guilty Actions Test allow for differentiating guilty participants from informed innocents? A re-examination. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *76*(1), 19–24. doi:10.1016/j.ijpsycho.2010.01.009

Gamer, M. (2011a). 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: Cambridge University Press.

Gamer, M. (2011b). Detecting deception using neuroimaging techniques. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 90–113). Cambridge: Cambridge University Press.

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*(12), 1287–1301. doi:10.1002/hbm.20343

Gamer, M., & Berti, S. (2010). Task relevance and recognition of concealed information have different influences on electrodermal activity and event-related brain potentials. *Psychophysiology*, *47*(2), 355–364. doi:10.1111/j.1469-8986.2009.00933.x

Gamer, M., & Berti, S. (2012). P300 amplitudes in the concealed information test are less affected by depth of processing than electrodermal responses. *Frontiers in Human Neuroscience*, *6*, 308. doi:10.3389/fnhum.2012.00308

Gamer, M., Gödert, 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: Official Journal of the International Organization of Psychophysiology*, *69*(1), 61–68. doi:10.1016/j.ijpsycho.2008.03.001

Gamer, M., Klimecki, O., Bauermann, T., Stoeter, P., & Vossel, G. (2012). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Social Cognitive and Affective Neuroscience*, *7*(5), 506–515. doi:10.1093/scan/nsp005

Gamer, M., Kosiol, D., & Vossel, G. (2010). Strength of memory encoding affects physiological responses in the Guilty Actions Test. *Biological Psychology*, *83*(2), 101–107. doi:10.1016/j.biopsycho.2009.11.005

Gamer, M., Rill, H.-G., Vossel, G., & Gödert, H. W. (2006). Psychophysiological and vocal measures in the detection of guilty knowledge. *International Journal of Psychophysiology*, *60*(1), 76–87. doi:10.1016/j.ijpsycho.2005.05.006

Gamer, M., Verschuere, B., Crombez, G., & Vossel, G. (2008). Combining physiological measures in the detection of concealed information. *Physiology & Behavior*, *95*(3), 333–340. doi:10.1016/j.physbeh.2008.06.011

Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *NeuroImage*, *55*(1), 312–319. doi:10.1016/j.neuroimage.2010.11.025

Giesen, M., & Rollison, M. A. (1980). Guilty knowledge versus innocent associations: Effects of trait anxiety and stimulus context on skin conductance. *Journal of Research in Personality*, *14*(1), 1–11. doi:10.1016/0092-6566(80)90035-5

Gläscher, J. (2009). Visualization of Group Inference Data in Functional Neuroimaging. *Neuroinformatics*, *7*(1), 73–82. doi:10.1007/s12021-008-9042-x

Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human Brain Mapping*, *12*(1), 1–19.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*(9), 394–400.

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases., *148*, 839–843.

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a Glance: Using Eye Movements to Investigate the Cognitive Neuroscience of Memory. *Frontiers in Human Neuroscience*, *4*. doi:10.3389/fnhum.2010.00166

Hira, S. (2003). The P300-based guilty knowledge test: Does it stand the test of time? *Psychophysiology*, 40, S10-S11.

Hira, S., & Furumitsu, I. (2009). Tonic arousal during field polygraph tests in guilty vs. innocent suspects in Japan. *Applied Psychophysiology and Biofeedback*, *34*(3), 173–176. doi:10.1007/s10484-009-9088-z

Honts, C. R. (2004). The psychophysiological detection of deception. In P. A. Granag & L. A. Strömwall (Eds.). *The Detection of Deception in Forensic Contexts* (pp. 103–123). Cambridge University Press, Cambridge.

Hu, X., Hegeman, D., Landry, E., & Rosenfeld, J. P. (2012). Increasing the number of irrelevant stimuli increases ability to detect countermeasures to the P300- based Complex Trial Protocol for concealed information detection. *Psychophysiology*, *49*(1), 85–95. doi:10.1111/j.1469-8986.2011.01286.x

Iacono, W. G., & Lykken, D. T. (1997). The validity of the lie detector: Two surveys of scientific opinion. *Journal of Applied Psychology*, *82*(3), 426–433. doi:10.1037/0021-9010.82.3.426

Iidaka, T., Matsumoto, A., Nogawa, J., Yamamoto, Y., & Sadato, N. (2006). Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cerebral Cortex*, *16*(9), 1349–1360. doi:10.1093/cercor/bhl040

Kensinger, E. A. (2009). Remembering the Details: Effects of Emotion. *Emotion Review*, *1*(2), 99–113. doi:10.1177/1754073908100432

Kiehl, K. A., Laurens, K. R., Duty, T. L., Forster, B. B., & Liddle, P. F. (2001). Neural sources involved in auditory target detection and novelty processing: an event-related fMRI study. *Psychophysiology*, *38*(1), 133–142.

Kim, J. S.-C., Vossel, G., & Gamer, M. (2013). Effects of Emotional Context on Memory for Details: The Role of Attention. *PLoS ONE*, *8*(10), e77405. doi:10.1371/journal.pone.0077405

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience*, *7*(1), 54–64. doi:10.1038/nrn1825

LaBar, K. S., & Phelps, E. A. (1998). Arousal-Mediated Memory Consolidation: Role of the Medial Temporal Lobe in Humans. *Psychological Science*, *9*(6), 490–493. doi:10.1111/1467-9280.00090

Labkovsky, E., & Rosenfeld, J. P. (2012). The P300-based, complex trial protocol for concealed information detection resists any number of sequential countermeasures against up to five irrelevant stimuli. *Applied Psychophysiology and Biofeedback*, *37*(1), 1–10. doi:10.1007/s10484-011-9171-0

Langleben, D. D., Schroeder, L., Maldjian, J. A., 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*(3), 727–732. doi:10.1006/nimg.2001.1003

Leal, S., & Vrij, A. (2010). The occurrence of eye blinks during a guilty knowledge test. *Psychology, Crime & Law*, *16*, 349–357. doi:10.1080/10683160902776843

Linden, D. E. J., Prvulovic, D., Formisano, E., Völlinger, M., Zanella, F. E., Goebel, R., & Dierks, T. (1999). The Functional Neuroanatomy of Target Detection: An fMRI Study of Visual and Auditory Oddball Tasks. *Cerebral Cortex*, *9*(8), 815 –823. doi:10.1093/cercor/9.8.815

Lorig, T. S. (2007). The respiratory system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson, (Eds.). *Handbook of Psychophysiology*, (pp. 231-244). Cambridge University Press.

Lubow, R. E., & Fein, O. (1996). Pupillary size in response to a visual guilty knowledge test: New technique for the detection of deception. *Journal of Experimental Psychology: Applied*, *2*(2), 164–177. doi:10.1037/1076-898X.2.2.164

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. *The American Psychologist*, *29*(10), 725–739.

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

Lynn, R. (1966). *Attention, Arousal and the Orienting Reaction*. New York: Pergamon.

MacLaren, V. V. (2001). A quantitative review of the guilty knowledge test. *The Journal of Applied Psychology*, *86*(4), 674–683. doi:10.1037//0021-9010.86.4.674

Maldjian, J. A., Laurienti, P. J., Burdette, J. B., & Kraft, R. A. (2003). An Automated Method for Neuroanatomic and Cytoarchitectonic Atlas-based Interrogation of fMRI Data Sets, *Neuroimage. 19*, 1233-1239.

Matsuda, I., Nittono, H., & Allen, J. J. B. (2012). The Current and Future Status of the Concealed Information Test for Field Use. *Frontiers in Psychology*, *3*. doi:10.3389/fpsyg.2012.00532

Matsuda, I., Nittono, H., Hirota, A., Ogawa, T., & Takasawa, N. (2009). Event-related brain potentials during the standard autonomic-based concealed information test. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *74*(1), 58–68. doi:10.1016/j.ijpsycho.2009.07.004

McNeil, B. J., & Hanley, J. A. (1983). Statistical approaches to the analysis of receiver operating characteristic (ROC) curves. *Medical Decision Making : An International Journal of the Society for Medical Decision Making*, *4*(2), 137–150.

Meijer, E., Verschuere, B., & Merckelbach, H. (2010). Detecting criminal intent with the Concealed Information Test. *The Open Criminology Journal*, *3*, 44–47.

Meixner, J. B., & Rosenfeld, J. P. (2011). A mock terrorism application of the P300-based concealed information test. *Psychophysiology*, *48*(2), 149–154. doi:10.1111/j.1469-8986.2010.01050.x

Mitchell, D. B. (2006). Nonconscious Priming After 17 Years Invulnerable Implicit Memory? *Psychological Science*, *17*(11), 925–929. doi:10.1111/j.1467-9280.2006.01805.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*(1), 49–100. doi:10.1006/cogp.1999.0734

Miyake, Y., Mizutanti, M., & Yamahura, T. (1993). Event Related Potentials as an Indicator of Detecting Information in Field Polygraph Examinations, *Polygraph*, 22. 131-149.

Nahari, G., & Ben-Shakhar, G. (2011). Psychophysiological and behavioral measures for detecting concealed information: The role of memory for crime details. *Psychophysiology*. *48*, 733-744. doi:10.1111/j.1469-8986.2010.01148.x

Nakayama, M. (2002). Practical use of the concealed information test for criminal investigation in Japan. In M. Kleiner (Ed.), *Handbook of polygraph testing* (pp. 49-86). San Diego, CS: Academic Press.

National Research Council (2003). *The Polygraph and Lie Detection. Committee to Review the Scientific Evidence on the Polygraph. Division of Behavioral and Social Sciences and Education.* Washington, DC: The National Academies Press.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, Affective & Behavioral Neuroscience*, *3*(1), 17–26.

Noordraven, E., & Verschuere, B. (2013). Predicting the Sensitivity of the Reaction Time-based Concealed Information Test. *Applied Cognitive Psychology*, *27*(3), 328–335. doi:10.1002/acf.2910

Nose, I., Murai, J., & Taira, M. (2009). Disclosing concealed information on the basis of cortical activations. *NeuroImage, 44*(4), 1380–1386. doi:10.1016/j.neuroimage.2008.11.002

Osugi, A. (2011). Daily application of the Concealed Information Test: Japan. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 253–175). Cambridge: Cambridge University Press.

Öhman, A., Hamm, A., & Hugdahl, K. (2000). Cognition and the autonomic nervous system: Orienting, anticipation, and conditioning. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology (2nd ed.)* (pp. 533–575). New York, NY, US: Cambridge University Press.

Patrick, C. J., & Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, *76*(2), 229–238. doi:10.1037/0021-9010.76.2.229

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, *48*(2), 175–187. doi:10.1016/j.neuron.2005.09.025

Phelps, E. A., & Sharot, T. (2008). How (and Why) Emotion Enhances the Subjective Sense of Recollection. *Current Directions in Psychological Science*, *17*(2), 147 –152. doi:10.1111/j.1467-8721.2008.00565.x

Podlesny, J. A., & Raskin, D. C. (1977). Physiological measures and the detection of deception. *Psychological Bulletin*, *84*(4), 782–799. doi:10.1037/0033-2909.84.4.782

Pollina, D. A., Dollins, A. B., Senter, S. M., Brown, T. E., Pavlidis, I., Levine, J. A., & Ryan, A. H. (2006). Facial skin surface temperature changes during a “concealed information” test. *Annals of Biomedical Engineering*, *34*(7), 1182–1189. doi:10.1007/s10439-006-9143-3

Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin*, *89*(3), 506–540. doi:10.1037/0033-2909.89.3.506

Raskin, D. C. (1989). Polygraph techniques for the detection of deception. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 247–296). New York, NY, US: Springer Publishing Co.

Reid, J. E., & Inbau, F. E. (1966). *Truth and Deception: The Polygraph ("Lie-detector") Technique*. William & Wilkins Company.

Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg & P. Hertel, (Eds.). *Memory and Emotion* (pp. 3-41). Oxford; New York: Oxford University Press.

Riggs, L., McQuiggan, D. A., Farb, N., Anderson, A. K., & Ryan, J. D. (2011). The role of overt attention in emotion-modulated memory. *Emotion*, *11*(4), 776–785. doi:10.1037/a0022591

Ritchey, M., Dolcos, F., & Cabeza, R. (2008). Role of Amygdala Connectivity in the Persistence of Emotional Memories Over Time: An Event-Related fMRI Investigation. *Cerebral Cortex, 18*(11), 2494 –2504. doi:10.1093/cercor/bhm262

Rosenfeld, J. P. (2011). P300 in detecting concealed information. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 63–89). Cambridge: Cambridge University Press.

Rosenfeld, J, P., Biroschak, J. R., & Furedy, J. J. (2006). P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *60*(3), 251–259. doi:10.1016/j.ijpsycho.2005.06.002

Rosenfeld, J. P., Cantwell, B., Nasman, V., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A Modified, Event-Related Potential-Based Guilty Knowledge Test. *International Journal of Neuroscience*, *42*(1), 157–161. doi:10.3109/00207458808985770

Rosenfeld, J. P., & Labkovsky, E. (2010). New P300-based protocol to detect concealed information: resistance to mental countermeasures against only half the irrelevant stimuli and a possible ERP indicator of countermeasures. *Psychophysiology*, *47*(6), 1002–1010. doi:10.1111/j.1469-8986.2010.01024.x

Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): a new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology*, *45*(6), 906–919. doi:10.1111/j.1469-8986.2008.00708.x

Rosenfeld, J. P., Shue, E., & Singer, E. (2007). Single versus multiple probe blocks of P300-based concealed information tests for self-referring versus incidentally obtained information. *Biological Psychology*, *74*(3), 396–404. doi:10.1016/j.biopsycho.2006.10.002

Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology, 41*(2), 205–219. doi:10.1111/j.1469-8986.2004.00158.x

Russ, M. O., Mack, W., Grama, C.-R., Lanfermann, H., & Knopf, M. (2003). Enactment effect in memory: evidence concerning the function of the supramarginal gyrus. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *149*(4), 497–504. doi:10.1007/s00221-003-1398-4

Ryan, J. D., Hannula, D. E., & Cohen, N. J. (2007). The obligatory effects of memory on eye movements. *Memory*, *15*(5), 508–525. doi:10.1080/09658210701391022

Schwedes, C., & Wentura, D. (2012). The revealing glance: Eye gaze behavior to concealed information. *Memory & Cognition*, *40*(4), 642–651. doi:10.3758/s13421-011-0173-1

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess “guilty knowledge.” *The Journal of Applied Psychology*, *85*(1), 30–37.

Seymour, T. L, & Fraynt, B. R. (2009). Time and encoding effects in the concealed knowledge test. *Applied Psychophysiology and Biofeedback*, *34*(3), 177–187. doi:10.1007/s10484-009-9092-3

Seymour, T. L, & Kerlin, J. R. (2008). Successful detection of verbal and visual concealed knowledge using an RT-based paradigm. *Applied Cognitive Psychology*, *22*(4), 475–490. doi:10.1002/acf.1375

Sharot, T., Verfaellie, M., & Yonelinas, A. P. (2007). How Emotion Strengthens the Recollective Experience: A Time-Dependent Hippocampal Process. *PLoS ONE*, *2*(10), e1068. doi:10.1371/journal.pone.0001068

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. doi:10.1016/j.cognition.2007.03.002

Siddle, D. A. T., O'Gorman, J. G., & Wood, L. (1979). Effects of Electrodermal Lability and Stimulus Significance on Electrodermal Response Amplitude to Stimulus Change. *Psychophysiology*, *16*(6), 520–527. doi:10.1111/j.1469-8986.1979.tb01514.x

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 592–604. doi:10.1037/0278-7393.4.6.592

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Oxford: Pergamon.

Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, *16*(4), 413–424. doi:10.1007/BF02352267

Steinmetz, K. R. M., & Kensinger, E. A. (2013). The emotion-induced memory trade-off: more than an effect of overt attention? *Memory & Cognition*, *41*(1), 69–81. doi:10.3758/s13421-012-0247-8

Stern, R. M., Breen, J. P., Watanabe, T., & Perry, B. S. (1981). Effect of feedback of physiological information on responses to innocent associations and guilty knowledge. *Journal of Applied Psychology*, *66*(6), 677–681. doi:10.1037/0021-9010.66.6.677

Suzuki, R., Nakayama, M., & Furedy, J. J. (2004). Specific and Reactive Sensitivities of Skin Resistance Response and Respiratory Apnea in a Japanese Concealed Information Test (CIT) of Criminal Guilt. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *36*(3), 202–209. doi:10.1037/h0087230

Timm, H. W. (1982). Analyzing deception from respiration patterns. *Journal of Police Science & Administration*, *10*(1), 47–51.

Twyman, N. W., Moffitt, K., Burgoon, J. K., & Marchak, F. (2010). Using Eye Tracking Technology as a Concealed Information Test. *Hawaii International Conference on*

System Sciences, *43*(Proceedings of the Credibility Assessment and Information Quality in Government and Business Symposium), 48–54.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain, *Neuroimage, 15(1)*, 273-89.

Velden, M., & Wölk, C. (1987). Depicting cardiac activity over real time: a proppposal for standardization. *Journal of Psychophysiology, 1*, 173–175.

Venables, P. H., & Christie, M. J. (1980). Elecrodermal activity. In I. Martin & P. H. Venables (Eds.), *Techniques in Psychophysiology* (pp. 3 – 67). Chichester, U.K.: Wiley.

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. 128–150). Cambridge: Cambridge University Press.

Verschuere, B., Crombez, G., De Clercq, A., & Koster, E. H. W. (2004). Autonomic and behavioral responding to concealed information: Differentiating orienting and defensive responses. *Psychophysiology*, *41*(3), 461–466. doi:10.1111/j.1469-8986.00167.x

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*(7), 991–1002. doi:10.1002/acf.1601

Verschuere, B., Crombez, G., Koster, E. H. W., & De Clercq, A. (2007). Antisociality, underarousal and the validity of the Concealed Information Polygraph Test. *Biological Psychology*, *74*(3), 309–318. doi:10.1016/j.biopsycho.2006.08.002

Verschuere, B., Crombez, G., Koster, E. H. W., Van Bockstaele, B., & De Clercq, A. (2007). Startling secrets: startle eye blink modulation by concealed crime information. *Biological Psychology*, *76*(1-2), 52–60. doi:10.1016/j.biopsycho.2007.06.001

Verschuere, B., Crombez, G., Smolders, L., & Clercq, A. D. (2009). Differentiating Orienting and Defensive Responses to Concealed Information: The Role of Verbalization.

Applied Psychophysiology and Biofeedback, *34*(3), 237–244. doi:10.1007/s10484-009-9093-2

Verschuere, B., & De Houwer, J. (2011). Detecting concealed infomration 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. 46–62). Cambridge: Cambridge University Press.

Wallin, B. G. (1981). Sympathetic Nerve Activity Underlying Electrodermal and Cardiovascular Reactions in Man. *Psychophysiology*, *18*(4), 470–476. doi:10.1111/j.1469-8986.1981.tb02483.x

Winograd, M. R., & Rosenfeld, J. P. (2011). Mock-crime application of the Complex Trial Protocol (CTP) P300-based concealed information test. *Psychophysiology*, *48*(2), 155–161. doi:10.1111/j.1469-8986.2010.01054.x

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 : Official Journal of the International Organization of Psychophysiology*, *84*(2), 140-148.

6. Abstract

The Concealed Information Test (CIT) is an implicit memory test, using multiple-choice questions to detect crime-related knowledge (Lykken, 1959, 1998). The single answer options presented after a question should be equally plausible and undistinguishable for an innocent person. Therefore, only a guilty person would recognize the correct answer option (i.e., the relevant detail) and accordingly show increased skin conductance responses, accompanied by decreased cardiovascular and respiration responses (Gamer et al, 2006). A huge amount of studies reported the CIT to detect concealed information with a high validity (Ben-Shakhar & Elaad, 2003). Nevertheless, the external validity of the CIT is still debated (Honts, 2004) and especially the influence of emotional factors on the memory for relevant details of a crime is unclear. Basic research in emotional memory reported an advantage for emotional arousing compared to neutral information (e.g., Kensinger, 2009). Moreover, emotional arousal was reported to increase memory for central details at expense of memory for peripheral details (cf., Christianson, 1992).

The current thesis aimed to investigate the CIT under laboratory conditions that stronger approximate real-life settings compared to previous studies (e.g., encoding of relevant details only incidentally during the mock-crime). In addition, the influence of emotional factors on the encoding of relevant details of a crime and their retrieval during the CIT were examined in further detail. Therefore, multiple approaches (i.e., autonomic response measures, eye-tracking, neuroimaging) were used to investigate these issues.

In the first study, the amount of emotional arousal during the mock-crime and the time of the CIT investigation were manipulated, as well as the type of relevant details (i.e., central versus peripheral details of the mock-crime). In addition to the traditionally used autonomic measures, ocular responses were recorded during a CIT that was conducted immediately after the mock-crime or two weeks later. The current results revealed that emotional arousal

strengthened the autonomic responses during a delayed CIT, when asking for central aspects of the mock-crime. Additionally, ocular measures were able to detect concealed information on central details in specific time intervals.

The second study of the current thesis investigated the influence of the encoding context (i.e., criminal versus non-criminal, enactment versus intention) on the retrieval of relevant details during a CIT using fMRI. Therefore, three groups were investigated with the same CIT: guilty subjects who committed a mock-crime (ACT), guilty subjects who only planned a mock-crime (PLAN), innocent subjects who encoded the relevant details in an non-criminal context (INNOCENT). In line with previous research (cf., Gamer, 2011b), a ventral fronto-parietal brain network was found to detect memory for known relevant details and no group differences in brain activation in these ROIs were found. Nevertheless, differences in neural activity between the ACT and PLAN group were found in brain regions reported to reflect the enactment of a task (i.e., the SMG). For the ACT and INNOCENT group, activation differences in regions reported to be relevant for enhanced emotional memory (i.e., amygdala and hippocampus) were found.

In sum, memory for relevant details was found to drive the physiological responses during the CIT. Independent of the encoding context, the recognition of relevant details during an fMRI-based CIT increased the activity in a ventral fronto-parietal brain network. However, increased emotional arousal during information encoding was found to modulate the autonomic responses during the CIT. This effect was especially pronounced for central crime details in a delayed CIT investigation. This finding further supports the external validity of the CIT. Overall, emotional arousal is not necessary for detection of concealed information in the CIT, but can increase the detection accuracy under specific conditions.

7. Zusammenfassung (German Summary)

Der Tatwissentest (TWT) ist ein impliziter Gedächtnistest, der Multiple-Choice Fragen nutzt um verheimlichtes Wissen aufzudecken (Lykken, 1959, 1998). Die einzelnen Antwortoptionen sind gleich wahrscheinlich für Unschuldige. Nur ein Täter sollte die korrekte Antwortalternative (das relevante Tatdetail) erkennen und einen Anstieg in der Hautleitreaktion sowie eine Verminderung der kardiovaskulären und respiratorischen Reaktion zeigen (Gamer et al., 2006). Viele Studien belegen die kriterienorientierte Validität Gültigkeit des TWT (Ben-Shakhar & Elaad, 2003), trotzdem wurde seine externe Validität diskutiert (Honts, 2004). Besondere der Einfluss emotionaler Faktoren ist in diesem Zusammenhang unklar. Grundlagenwissenschaftliche Studien berichten einen Vorteil für emotionale im Vergleich zu neutralen Gedächtnisinhalten (vgl., Kensinger, 2009). Zudem scheint emotionale Erregung die Erinnerung für zentrale Details zu verbessern und für periphere Details zu verschlechtern (vgl., Christianson, 1992).

Die vorliegende Doktorarbeit untersuchte den TWT unter Laborbedingungen, die versuchten realistische Untersuchungsumstände besser abzubilden (z.B., war die Enkodierung der relevanten Details nur während des Scheinverbrechens möglich). Zudem sollte der Einfluss emotionaler Faktoren während der Informationsenkodierung und dem Abruf im TWT genauer untersucht werden. Dabei wurden verschiedene methodische Ansätze verwendet (z.B., autonome Reaktionsmaße, Augenbewegungsaufzeichnungen, neuronale Bildgebung).

In der ersten Studie wurden das Ausmaß an emotionaler Erregung während des Scheinverbrechens, die Zeitdauer bis zum TWT und die Art des relevanten Details (zentral oder peripher) manipuliert. Zusätzlich zu den traditionellen, autonomen Reaktionsmaßen wurden Augenbewegungen während des TWT aufgezeichnet. Der Test wurde entweder direkt im Anschluss an das Scheinverbrechen durchgeführt oder zwei Wochen später. Die vorliegenden Ergebnisse zeigen, dass emotionale Erregung während des Scheinverbrechens die autonomen Reaktionen im TWT verstärken kann, besonders wenn in einem zeitlich verzögerten TWT nach zentralen Tatdetails gefragt wird. Zusätzlich wurde gefunden, dass Augenbewegungen in konkreten Zeitfenstern verheimlichtes Wissen über zentrale Tatdetails aufdecken können.

Die zweite Studie untersuchte den Einfluss des Enkodierungskontexts auf den Abruf relevanter Tatdetails im TWT mittels fMRT. Hierfür wurden drei Gruppen mit demselben TWT untersucht: Täter die ein Scheinverbrechen ausübten (ACT), Täter die ein Scheinverbrechen nur planten (PLAN), informierte Unschuldige die Tatdetails in einem nicht kriminellen Kontext kennenlernten (INNOCENT). Entsprechend vorheriger Forschungsergebnisse (vgl., Gamer, 2011b), war ein Netzwerk aus ventralen fronto-parietalen Hirnregionen in der Lage ist die Erinnerung an bekannte relevante Details aufzudecken. Es wurden keine Unterschiede in diesen Arealen zwischen den Gruppen gefunden. Gruppenunterschiede zeigen die ACT und PLAN Gruppe im SMG, einer Hirnregion die bei Handlungsausführungen beteiligt ist. ACT und INNOCENT Gruppe unterschieden sich in der Aktivierung von Amygdala und Hippocampus, die für verbesserte Erinnerung emotionaler Informationen relevant sind.

Es wurde gezeigt, dass die Erinnerung an relevante Tatdetails die physiologischen Reaktionen während des TWT steuert. Unabhängig vom Enkodierungskontext, führte die Wiedererkennung relevanter Tatdetails in einem TWT zu erhöhten Aktivierungen in einem ventralen fronto-parietalen Hirnnetzwerk. Allerdings wurde auch gezeigt, dass eine erhöhte emotionale Erregung während des Scheinverbrechens die autonomen Reaktionsmaße während des TWT stärkt. Dieser Einfluss war besonders ausgeprägt für die Wiedererkennung zentraler Tatdetails in einem zeitlich verzögerten TWT. Dieser Befund unterstützt die externe Validität des TWT. Demnach scheint emotionale Erregung keine notwendige Bedingung für den TWT zu sein, aber kann die korrekte Aufdeckung verheimlichter Informationen unter bestimmten Bedingungen verbessern.

8. Appendix

Publications

Study I of the current thesis refers to the following publications:

Peth, J., Vossel, G., & Gamer, M. (2012). Emotional arousal modulates the encoding of crime-related details and corresponding physiological responses in the Concealed Information Test. *Psychophysiology*, *49*(3), 381–390. doi:10.1111/j.1469-8986.2011.01313.x

Peth, J., Kim, J. S.-C., & Gamer, M. (2013). Fixations and eye-blinks allow for detecting concealed crime related memories. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *88*(1), 96–103. doi:10.1016/j.ijpsycho.2013.03.003

Eidesstattliche Erklärungen

Eidesstattliche Erklärung nach § 8, Abs. 4, der Promotionsordnung der Medizinischen Fakultät der Universität Hamburg für Nicht-Mediziner und Nicht-Medizinerinnen zur Erlangung des akademischen Doktogrades PhD (alternativ Dr. rer. biol. Hum.) vom 19. Oktober 2011.

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Andere als die angegebenen Quellen und Hilfsmittel habe ich nicht benutzt und die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht. Zudem erkläre ich hiermit, dass die von mir vorgelegte Dissertation nicht Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

Hamburg, den ______________________ ________________________________

Unterschrift

Acknowledgement

The current thesis was written at the Department of Systems Neuroscience at the University Medical Center Hamburg-Eppendorf and I would like to thank Dr. Matthias Gamer and Prof. Christian Büchel for this opportunity. I would especially like to thank my advisor, Dr. Matthias Gamer for his continuous support and guidance throughout all phases of this dissertation. His constructive feedback was always helpful and I really appreciate his scientific input.

I would like to thank Prof. Christoph Mulert and Prof. Jan Wacker for agreeing to be members of the dissertation committee.

Moreover, I greatly appreciate the support by my friends and colleagues at the Department of Systems Neuroscience and would especially like to thank Janine Bayer, Sabrina Boll and Kristina Suchotzki for thoroughly proofreading this manuscript and giving helpful feedback.

I thank Michael Gründler, Hannes Mehrer, Hellena Salostowitz and Philine Böttcher for their support during data collection.

Last but not least, I would like to thank my friends (in Berlin & Hamburg) and my family for their support.