Running head: SEPARATING DECEPTION AND ORIENTING

Separating deceptive and orienting components in a Concealed Information Test

Wolfgang Ambach *, Rudolf Stark **, Martin Peper ***, Dieter Vaitl */**

* Institute for Frontier Areas of Psychology and Mental Health, Freiburg, Germany
** Bender Institute of Neuroimaging, University of Giessen, Germany
*** Department of Psychology, University of Freiburg and Department of Forensic Medicine, Forensic Psychiatric Services, University of Bern, Switzerland

Text pages: 43
Figures: 5
Tables: 3

Correspondence: Wolfgang Ambach
Institut für Grenzgebiete der Psychologie und Psychohygiene, Freiburg, Germany
Wilhelmstr. 3A
D-79098 Freiburg
Germany
Telephone: +49-(0)761-20721-25
Fax: +49-(0)761-20721-99
E-mail: ambach@igpp.de

Abstract

The Concealed Information Test (CIT) requires the examinee to deceptively deny recognition of known stimuli and to truthfully deny recognition of unknown stimuli. Because deception and orienting are typically coupled, it is unclear how exactly these sub-processes affect the physiological responses measured in the CIT. The present study aimed at separating the effects of deception from those of orienting. In a mock crime study, using a modified CIT, thirty-six of seventy-two subjects answered truthfully ('truth group'), whereas the other thirty-six concealed their knowledge ('lie group'). Answering was delayed for four seconds after item presentation. Electrodermal activity (EDA), respiration (RLL), and phasic heart rate (HR) were recorded. A decomposition of EDA responses revealed two response components; the response in the first interval was expected to indicate orienting, stimulus evaluation, and answer preparation, whereas the response in the second interval was assumed to reflect answer-related processes. Inconclusively, both EDA components differentiated between 'probe' and 'irrelevant' items in both groups. Phasic HR and RLL differed between item classes only in the 'lie' group, thus reflecting answer-related processes, possibly deception, rather than merely orienting responses. The findings further support the notion that psychophysiological measures elicited by a modified CIT may reflect different mental processes involved in orienting and deception.

Key Words

Concealed Information Test; Deception; Orienting reflex; Perception; Differentiation of deception.

Introduction

Deception and orienting during a Concealed Information Test

The well-known term 'lie detection' suggests that the investigator should be able to detect lying by identifying a lie-specific physiological response pattern. However, the *assumption that there is a distinctive pattern of physiological responses which accompanies lying and which can be distinguished from that which accompanies truth telling* did not find support in early lie-detection studies. Lying could only be detected when associated with other specific mental processes such as the recognition of known objects (Lykken, 1959). Thus, in the Concealed Information Test, subjects deceptively denying their deed-related knowledge could be differentiated from subjects truthfully denying having that knowledge. Using the CIT, a fixed coupling between the fact of lying and the existence of a specific knowledge was implemented in the research on lie detection and its application. For the purpose of investigating subjects accused of having committed a crime, the CIT paradigm, meanwhile improved over decades, accomplished relatively high correct-classification rates (Ben-Shakhar and Elaad, 2003), and, thus, it is of considerable forensic interest. For the application in a criminal context, it is of minor importance whether it is the specific, deed-related knowledge, or the deceptive behavior associated with its concealment, which is being detected in the guilty subject. Therefore, research in this field has mostly been motivated by the aim to improve the detection of concealed information, whereas deception per se was less often the focus.

The differentiation-of-deception paradigm

Furedy et al. (1988) emphasized a psychophysiological rather than a forensic view of deception when they introduced the *differentiation-of-deception paradigm* (DDP). This paradigm is based on the idea that deception should be studied in experiments that expose subjects to *two conditions (experimental and control) which differ only with respect to deception*. Instead of merely detecting guilt in individuals, deception as a psychological process with several involved sub-processes became the focus of interest. Subsequent studies aimed at overcoming the fixed coupling of deception with orienting to guilty knowledge and other mental processes and intended to disentangle potentially confounding factors (Furedy et al., 1991; Furedy et al., 1994; Vincent and Furedy, 1992). For this purpose, modifications of the paradigm have been applied, which comprise delay in answering, variations of response type or of contextual information, repetition of questioning blocks, or variations of the dependent measures.

Delay in answering

The idea of separating the intention to deceive from the act of deception by using a delay in the subjects' answers has been introduced by Dawson (1980) who employed a laboratory variant of the Control Question Test. Furedy et al. (1988, 1991) demanded partially immediate, partially delayed verbal answers from the subjects and concluded that this variation did not help differentiate between subjects answering deceptively and those answering honestly. In the same vein, Furedy and Ben-Shakhar (1991), who introduced a delayed-answer condition in a CIT, found that differential EDA response to known and unknown objects, measured immediately after question presentation, was unaffected by delayed answering and that the differential EDA response measured after the delayed answer was *markedly attenuated*. They found high differential EDA responses between item classes for the responses measured immediately after stimulus presentation, for all conditions with immediate, delayed, or missing answers. These differential responses have been discussed to be due to a stronger orienting to objects with greater personal significance, but also as being influenced by the intention to deceive (Furedy and Ben-Shakhar, 1991).

Deceptive vs. truthful answers

The role of lying behavior per se, in the psychophysiological detection of deception was first investigated with card test studies (Kugelmass et al., 1967). The finding that the requirement of a verbal answer increased detection rates was later interpreted in terms of the orienting reflex (e.g. Ben-Shakhar, 1977). The finding by Horneman and O'Gorman (1985) however, that "no" answers led to higher detection rates than "yes" answers, is not explained by the orienting theory.

A meta-analysis on the Guilty Knowledge Test revealed that a deceptive verbal response tended to increase test accuracy, but overall, this effect was not significant (Ben-Shakhar and Elaad, 2003).

Most of the previous CIT studies differentiated between 'guilty' and 'innocent' subjects; some authors used a variation in response type between 'always NO', 'always YES' and 'silent' (e.g. Furedy and Ben-Shakhar, 1991); however, there were only few studies wherein all subjects had the same 'guilty' knowledge and wherein subjects were divided into groups that differed only with respect to the instruction to answer truthfully vs. deceptively. Differentiation between deceptive and truthful answers can be made within or between subjects. Whereas Furedy et al. (1988), and Gödert et al. (2001), advised subjects to answer half of the questions deceptively and the other half honestly, according to a question-wise indication, Furedy and Ben-Shakhar (1991) studied the variation in the type of verbal answer between subjects.

Results consistently underpin that, although deceptive answering is not a necessary condition for the detection of concealed information, it enhances differential physiological response to *probe* vs. that to *irrelevant* items.

Dependent measures

In previous CIT studies, EDA was mostly the only dependent measure. However, phasic heart rate (HR) and respiration line length (RLL) are also known to be valid indicators of concealed information (Bradley and Janisse, 1981; Bradley and Rettinger, 1992); furthermore, heart rate has been shown to be sensitive to deceptive answering in the context of the CIT (Gödert et al., 2001). The issue of what these measures reflect in detail, with regard to the CIT as well as the psychophysiology of cognitive processes in general, is still under discussion. There is some evidence that the reflected processes differ at least partly from those reflected by EDA (Wölk et al., 1989; Barry and Maltzman, 1985; Barry, 1996; Verschuere et al., 2004; Gödert et al., 2006). To our knowledge, no previous CIT study contrasted the effects of truthful vs. deceptive answers on a combination of the measures EDA, pHR, and RLL.

Aim of the present study

The present mock-crime study aimed at differentiating the influences of the deceptive processes and the orienting processes on electrodermal, cardiovascular, and respiratory responses in a CIT. First, in the style of the differentiation-of-deception paradigm, but in a between-subjects design, subjects were advised to either answer truthfully ('truth group') or hide their knowledge deceptively ('lie group'), with all other experimental conditions being kept constant. Second, a splitting of sub-processes into separate time windows was accomplished by introducing a time delay of four seconds between the object/question presentation and the prompt to answer.

In detail, the expectations were the following:

1. In the CIT study, which involved delay in answering, subjects' physiological responses to the item presentations and the answer requests were expected to be temporally distinct. The first response is supposed to reflect orienting, stimulus evaluation, intentional factors, and answer preparation, whereas the second response should be associated with the act of (truthful or deceptive) responding.

2. The first physiological response component 'orienting' should therefore differ between recognized (known) and unrecognized (unknown) items in the 'lie' and the 'truth' group. If one further assumes that an item is more significant for a subject when it is associated with the instruction to deceive, or that the intention or preparation to deceive enhances physiological responses, then the differential response between both item classes should be larger in the 'lie' than in the 'truth' group.

3. The second physiological response, which is supposed to be independent from orienting, but related to processes of truthful or deceptive answering, should differ between known and unknown items in only the 'lie' group. A difference between groups should only be found in the response to the known but not to the unknown items.

4. From the different response patterns of EDA, phasic HR, and RLL, the degree to which each measure reflects either of the mentioned sub-processes can be inferred. In this line of reasoning, habituation effects of each measure in the different conditions are also of interest.

Materials and Methods

Subjects

Seventy-two healthy students of various faculties (24 m, 48 f; mean age 24.6 ± 3.6 years; 68 right-handed) voluntarily participated in the study. They were paid 10 Euros, with an additional incentive of 3 Euros. Informed consent was obtained from all subjects.

Mock-crime scenario

In a shielded room, subjects had to handle seven objects according to a task list. No criminal associations were suggested. They took one object after the other into their hands, looked at it in detail, and deposited it on top of a storage rack. Then, they packed all seven objects attentively from there into a suitcase placed in the same room. For each participant, one object had been randomly drawn from each of the seven object categories. The object categories, each comprising five objects, were household objects, key pendants, boxes, office material, cosmetics, wooden toy fruits, and drink packages. Subjects finally came across a small box which contained three Euros. They handed the money together with the filled suitcase to the experimenter who kept it until the end of the experiment. Participants were told that they would receive the money at the end of the experiment, provided they performed well in the interrogation task they would face in the second part of the experiment. Because participants were given different instructions in this later task, they were not informed about details of the task, and no specification of 'performing well' was made.

Concealed Information Test

A variant of the CIT, which consisted of seven *scenario-related* item categories, each comprising five questions with item presentations, was used. Questions and item pictures were presented foveally on a 19'' monitor at a distance of 90 cm for 10 seconds, followed by a blank screen for equally distributed 5.0–7.5 second intervals. Picture size was 6.0° by 8.0° of visual angle.

Seven *scenario-related* questions (e.g. 'Did you see this cosmetic in the deed room?') referred to the seven objects in the mock-crime study. Each of the *scenario-related probe* items was combined with four *scenario-related irrelevant* items which belonged to the same item category (e.g. key pendants, cosmetics); these had some categorical similarity, but were all unknown to the participants.

Preceding each category, two *neutral* items were presented as distractors. The *neutral* questions referred to everyday objects that had to be identified (e.g. 'Is this a pineapple?'). One of the two questions preceding each category had to be correctly answered with 'yes' and the other with 'no' (with a pseudorandomized sequence). Responses to these *neutral* questions were not evaluated.

Each *scenario-related* question was presented five times in sequence together with a different picture of the corresponding category, which was simultaneously displayed below it. Together with the two *neutral* questions preceding each category, this resulted in a total of 49 item presentations per run; the total experiment consisted of two runs. In order to reduce sequential effects, the two runs had different sequences of categories and different orders of items within each category.

Subjects were assigned randomly and without the knowledge of the experimenters to one of the two groups with different instructions. Half of them, assigned to the 'lie' group (12 m, 24 f), were instructed to deny their knowledge about things they had seen and touched in the scenario room, but to answer all other presented questions and items correctly. These participants had to say 'no' to all – *probe* and *irrelevant* – item presentations of the *scenario-related* type as well as to half of the *neutral* items, and they had to answer with 'yes' to the other half of the *neutral* questions. The other half of the subjects, assigned to the 'truth' group (12 m, 24 f), were instructed to answer truthfully, i.e. *as correct as possible*, to all questions.

Two indication fields containing question marks appeared with a delay of four seconds after a question was asked; this prompted the subjects to answer. Then, answers had to be given as quickly as possible both by pressing one of the two response keys and by vocally responding with 'yes' or 'no'. Key assignment was balanced across subjects. Following the answer, the given 'yes' or 'no' replaced the question marks and remained visible on the screen as long as the item question was presented (see figure 1).

*** insert figure 1 about here ***

Procedure

In order to help subjects make a clear distinction between the mock crime and the CIT, the two parts of the experiment were guided by two experimenters. After subjects had given their informed consent, they were led to the experimental room by the first experimenter; there, they were informed about the subsequent mock-crime scenario by a second experimenter. After the mock-crime procedure, subjects were led back to the first experimenter who performed the 'psychophysiological investigation'. They were connected to the polygraph leads and were asked to answer the CIT questions. The CIT was then initiated; the two main runs were preceded by a training run consisting of two blocks with five *neutral* items each. After completing both main runs of the CIT and disconnecting from leads, a memory test was performed on the subjects. In this test, all five pictures of each category, one item category after the other, were presented on the screen simultaneously, and subjects were asked to identify the probe item within each category. Finally, subjects received payment and the incentive of 3 Euros, regardless of their answers in the CIT.

Physiological measurement

The physiological recordings took place in a dimly lit, electrically and acoustically shielded experimental chamber (*Industrial Acoustics GmbH*, Niederkrüchten, Germany). Subjects sat in an upright position so that they could comfortably see the monitor and reach the keyboard. Temperature in the cabin was set at 21°C at the beginning of the first run, with an increase of maximum 2°C throughout the course of the experiment.

Skin conductance (EDA), respiratory activity, electrocardiogram (ECG), and finger plethysmogram were registered. Physiological measures were A/D-converted and logged by the *Physiological Data System I 410-BCS* manufactured by *J&J engineering (Poulsbo, Washington).* The A/D-converting resolution was 14 bit, allowing skin conductance to be measured with a resolution of 0.01 µS. All data were sampled with 510 Hz. Triggers indicating question onsets were registered with the same sampling frequency.

For skin-conductance recordings, standard Ag/AgCl electrodes (*Hellige;* diameter 0.8 cm), isotonic signa electrode creme (*Parker Laboratories Inc.*) and a constant voltage of 0.5 volts were used. The electrodes were fixed at thenar and hypothenar sites of the nondominant hand.

For registration of the respiratory activity, a PS-2 biofeedback respiration sensor belt (*KarmaMatters,* Berkeley, California) with a built-in length-dependent electrical resistance was used. It was fixed at the level of the lower thoracic aperture.

The ECG was measured with *Hellige* electrodes (diameter 1.3 cm) according to Einthoven II.

Finger pulse signal was transmitted by an infrared system in a cuff around the ring finger of the nondominant hand.

Behavioral measures

Subjects responded with 'yes' or 'no' by key presses as well as verbal expressions (the latter were not further analyzed). Key presses indicating 'yes' or 'no' answers were time-logged and stored on the stimulus-presenting computer for later evaluation of reaction times and error rates. Behavioral data were synchronized with physiological measures with an accuracy of ± 2 ms.

Data reduction and statistical analysis

Skin conductance reactions were assessed by a computerized method based on decomposition of overlapping reactions as proposed by Lim et al. (1997). This method was chosen because, due to the delay of four seconds between question and prompt to answer, two subsequent physiological reactions occurred with a short delay. As Lim (1999) showed in an experiment with short interstimulus intervals, conventional trough-to-peak evaluation is inadequate in this case, because the first of two reactions causes a diminishing bias in the estimation of the second one. The size of this bias is determined by the size of the first reaction and by the time interval between both reactions. Decomposition aims at overcoming this problem of overlapping EDA reactions.

First, in a curve-fitting procedure guided by minimizing error squares, model coefficients were optimized for each subject. Second, all trials were evaluated by decomposing EDA by use of the subject's individual model coefficients. The use of individually constant templates for the calculation of EDA responses reflects the assumption of individual response characteristics of the electrodermal system. Third, magnitudes of all EDA responses that were elicited within a time window of 0.5 to 4.5 seconds after item presentation were additively combined to a 'first response' (EDA_1). The sum of EDA responses which began between 4.5 and 8.5 seconds after item presentation, i.e. between 0.5 and 4.5 seconds after the subjects were prompted to answer, was calculated as 'second response' (EDA_2). As the decomposed responses were not calculated as trough-to-peak amplitudes but rather as relative activity, the unit of these measures was arbitrary and depended on the subject's individual template. (Transformation into µS was not required because a within-subject standardization followed.)

Respiratory data of four subjects had to be discarded from evaluation because of artifacts. After low-pass filtering, the total respiration line length (RLL) was automatically computed over a time interval of 10 seconds after trial onset. The RLL measure integrates information about frequency and depth of respiration. The method was derived from Timm (1982b) and modified by Kircher and Raskin (2003). ECG data obtained from one subject had to be excluded from analysis because of frequent extrasystoles. After notch filtering at 50 Hz, R-wave peaks were automatically detected and visually controlled. The R-R intervals were transformed into heart rate (HR) and real-time scaled (Velden and Wölk, 1987). The HR during the last second before trial onset served as pre-stimulus baseline. The phasic heart rate (pHR) was calculated by subtracting this value from each second-per-second poststimulus value. For extracting the trial-wise information of the phasic HR, the mean change in HR within 15 seconds after trial onset, compared with the pre-stimulus baseline, was calculated (Bradley and Janisse, 1981; Verschuere et al., 2007).

The finger pulse waveform length (FPWL) within the first 10 seconds after trial onset was calculated from the finger pulse waveform, and then subjected to further analyses (Elaad and Ben-Shakhar, 2006). The FPWL comprises information about both HR and pulse amplitude.

The delay between the prompt to answer, indicated by the appearance of the fields with question marks on the screen, and the pressing of the key by the subject was calculated as reaction time.

Standardization: A within-subject standardization of measured values has been proposed by Lykken and Venables (1971). Here, according to Ben-Shakhar (1985), Gronau et al. (2005), and Gamer et al. (2006), the physiological and behavioral measures are z-transformed for each subject and for each data channel. With respect to the criticism on a category-wise standardization (Meijer et al., 2007), all trials, including both runs of the experiment (*probe* and *irrelevant* trials of the *scenario-related* question type, not including the first trials of each stimulus category), were used to calculate individual means and standard deviations. The z-transformed values were used in subsequent statistical analyses.

Statistics: Statistical analyses were performed with *SPSS*, Version 12.0 (*SPSS Inc.*, Chicago).

For each physiological and behavioral measure, a 2*2*2 ANOVA was calculated for the data obtained from *scenario-related* trials, with the within-subject factors Probe (*probe* vs. *irrelevant* trials) and Run (first vs. second run), and the between-subjects factor Group ('truth' vs. 'lie' group). Since ANOVAs were calculated on the basis of within-subject z-values, no main effects for the factor Group are reported[1]. Significance level for the assessment of main and interaction effects was set to 0.05. Follow-up t-tests for matched samples (one-tailed, significance level 0.05) were carried out when interactions were found. Cohen's d was calculated as estimate of effect size (Cohen, 1988; Rosnow and Rosenthal, 1996).

In addition, the validity of each data channel and the validity of an optimized combination of measures were evaluated using a binary logistic regression analysis. This procedure aimed at identifying the contribution of the various measures to differentiate between *probe* and *irrelevant* items. The probability with which a trial has to be regarded as a *probe* trial is a function of the optimized linear combination of measures. Thus, trial type served as dependent variable and the various measures as predictors. In a first calculation, a stepwise inclusion of parameters was performed following a Wald statistic with probabilities for inclusion and exclusion set to 0.05 and 0.10, respectively. Second, in order to estimate the predictive value of each measure per se, they were included singularly in an analogous regression analysis. The logistic regression model was based on randomly selected 70% of trials and the remaining 30% served as cross-validation in order to estimate the shrinkage of the regression model.

The ROC curve reflects the item-differentiation capability of the single data channels and their combinations for all possible cutoff-points; the area under the ROC curve, varying between 0 and 1 with a chance level of 0.5, serves as an overall index of detection accuracy (Bamber, 1975; Ben-Shakhar and Elaad, 2003; Gronau et al., 2005). Note that the ROC data in the present study refer to the classification of single items as *probe* or *irrelevant*, regardless of their grouping in categories. Thus, the present data are not comparable with ROC data from CIT studies, which refer to a different data aggregation level, mainly the differentiation between '*guilty*' and '*innocent*' subjects.

Results

Table 1 summarizes means and standard errors of means of raw scores for each data channel, collapsed over both runs of the experiment.

*** insert table 1 about here ***

Figure 2 illustrates differential response to *probe* vs. *irrelevant* items for the 'truth' and the 'lie' group. Estimated effect sizes are depicted for each of the physiological and behavioral measures.

*** insert figure 2 about here ***

Behavioral data

2.6% of all probe and irrelevant items were followed by incorrect answers; these trials were discarded from further analyses.

In the concluding memory test, 99.8% of *probe* items were identified correctly.

The mean reaction times to the CIT questions were 745 ms with a standard deviation of 403 ms. ANOVA for reaction times did not yield a main effect for Probe ($F_{1,70}$ = 0.12) but revealed a Probe by Group interaction ($F_{1,70}$ = 33.50; $p < 0.001$). In the 'truth' group, mean reaction times to the CIT questions were shorter in *probe* than in *irrelevant* trials (T_{34} = -4.39; $p < 0.001$; d = -1.32), whereas in the 'lie' group, the effect was in the opposite direction (T_{36} = 4.14; $p < 0.001$; d = 1.22). A main effect for Run ($F_{1,70}$ = 10.05; $p < 0.01$) reflected faster reactions in the second compared to the first run; neither of the interactions Probe by Run ($F_{1,70}$ = 2.66), Run by Group ($F_{1,70}$ = 0.71) or Run by Probe by Group ($F_{1,70}$ = 3.08) was significant.

Skin conductance

Figure 3 shows the averaged intra-trial course of skin conductance depicting grand means for *probe* and *irrelevant* items separately for both experimental groups.

*** insert figure 3 about here ***

Grand means show two EDA response components with an onset and peak asynchrony of four seconds.

After decomposition of skin conductance reactions and z-standardization for each component, ANOVA for EDA_1 showed a main effect for Probe ($F_{1,70} = 77.16$; $p < 0.001$) and a Probe by Group interaction ($F_{1,70} = 6.63$; $p < 0.05$). Reactions to *probe* items were greater than to *irrelevant* items in the 'truth' and the 'lie' group; this effect was greater in the 'lie' group ($T_{34} = 5.27$; $p < 0.001$; d = 1.58 and $T_{36} = 7.37$; $p < 0.001$; d = 2.12, respectively). A main effect for Run ($F_{1,70} = 38.92$; $p < 0.001$) indicated overall habituation of this component; the interaction Probe by Run ($F_{1,70} = 13.31$; $p < 0.01$) reflected a decrease in differential responsivity from the first to the second run. Neither of the interactions Run by Group ($F_{1,70} = 0.36$) and Run by Probe by Group ($F_{1,70} = 0.03$) was significant.

ANOVA for EDA_2 yielded a main effect for Probe ($F_{1,70} = 33.49$; $p < 0.001$) and a Probe by Group interaction ($F_{1,70} = 7.48$; $p < 0.01$). Greater responses to *probe* compared with *irrelevant* items were found in both groups; the greater effect size was found in the 'lie' group ($T_{36} = 6.04$; $p < 0.001$; d = 1.76 for the 'lie' group and $T_{34} = 2.16$; $p < 0.05$; d = 0.65 for the 'truth' group). A main effect for Run ($F_{1,70} = 19.28$; $p < 0.001$) indicated overall habituation; the interaction Probe by Run ($F_{1,70} = 6.86$; $p <$

0.05) reflected a decrease in differential responsivity. Neither of the interactions Run by Group ($F_{1,70} = 2.12$) and Run by Probe by Group ($F_{1,70} = 0.79$) was significant. An additional ANOVA was conducted in order to assess differential validity of both EDA components. The analysis comprised both EDA components and the additional factor 'Component', which differentiated between EDA_1 and EDA_2. Results show a marginally significant interaction of Component by Run ($F_{1,70} = 4.02$; $p < 0.05$), which reflects a lower habituation of EDA_2 from the first to the second run. However, the interactions *Component by Group* and *Component by Probe by Group* were tested negative ($F_{1,70} = 0.23$; $p > 0.05$, and $F_{1,70} = 0.19$; $p > 0.05$, respectively), so that the instruction - to deceive or not - is not shown to differentially affect the two EDA components.

Respiration

ANOVA for RLL data showed a main effect for Probe ($F_{1,66} = 33.91$; $p < 0.001$) and a Probe by Group interaction ($F_{1,66} = 18.33$; $p < 0.001$). In the 'lie' group, respiratory activity was significantly lower in *probe* than in *irrelevant* trials ($T_{34} = -5.94$; $p < 0.001$; $d = -1.76$), but no statistically significant differences between trial types were found in the 'truth' group ($T_{32} = -1.55$; $p > 0.1$). No sign of habituation was found; no main effect for Run ($F_{1,66} = 2.54$) was found; neither of the interactions Probe by Run ($F_{1,66} = 2.93$), Run by Group ($F_{1,66} = 2.66$) or Run by Probe by Group ($F_{1,66} = 0.58$) was significant.

Heart rate

Second-per-second values of HR after trial onset for *probe* and *irrelevant* items are depicted separately for both experimental groups in figure 4.

In the 'truth' group, two phasic HR increases were seen beginning about one second and about five seconds after trial onset, whereby *probe* and *irrelevant* trials showed a similar pattern. In contrast, in the 'lie' group, the HR showed a marked decrease in *probe* trials compared to an increase in *irrelevant* trials.

ANOVA for pHR data, averaged over fifteen seconds after trial onset, showed a main effect for Probe ($F_{1,69} = 49.75$; $p < 0.001$) and a Probe by Group interaction ($F_{1,69} = 34.71$; $p < 0.001$). Enhanced HR decelerations after *probe* items compared with *irrelevant* items were confirmed for the 'lie' but not for the 'truth' group ($T_{36} = -8.84$; $p < 0.001$; d = -2.59 and $T_{33} = -0.86$; p > 0.1, respectively). No sign of habituation was found; a main effect for Run ($F_{1,69} = 3.51$) was not found; neither of the interactions Probe by Run ($F_{1,69} = 1.40$), Run by Group ($F_{1,69} = 0.85$) or Run by Probe by Group ($F_{1,69} = 1.19$) was significant.

*** insert figure 4 about here ***

Finger pulse

ANOVA for FPWL indicated a main effect for Probe ($F_{1,70} = 31.05$; $p < 0.001$) and a Probe by Group interaction ($F_{1,70} = 10.11$; $p < 0.01$). FPWL values were lower after *probe* items than after *irrelevant* items in the 'truth' and the 'lie' group; also this effect was more pronounced in the 'lie' group ($T_{34} = -2.25$; $p < 0.05$; d = -0.68 and $T_{36} = -5.28$; $p < 0.001$; d = -1.50, respectively). No main effect for Run ($F_{1,70} = 1.10$) was found; the interaction Probe by Run ($F_{1,70} = 7.69$; $p < 0.01$) reflected a decrease in differential responsivity from the first to the second run. Neither of the interactions Run by Group ($F_{1,70} = 1.61$) and Run by Probe by Group ($F_{1,70} = 0.10$) was significant.

Logistic regression model

The binary logistic regression and the following results refer to a discrimination between *probe* and *irrelevant* trials on a single-trial basis. For the 'lie' group, the predictors EDA_1, EDA_2, pHR, RLL, RT, and FPWL were included successively following the Wald statistic; for the 'truth' group, EDA_1, EDA_2, RLL, and RT were included successively. Optimal weights, standard errors, Wald statistic, and significance level are summarized for both experimental groups in table 2.

*** insert table 2 about here ***

Rate of correct trial type classification (criterion: $p > 0.25$ for classification as *probe*) was 73.0% for the 'lie' group (cross-validation: 73.1%) and 63.2% for the 'truth' group (cross-validation: 62.3%).

In addition to the stepwise inclusion, analogous analyses were performed with singular inclusion of each of the measures: 'first electrodermal response' (EDA_1), 'second electrodermal response' (EDA_2), ‘phasic heart rate’ (pHR), 'respiration line length’ (RLL), ‘finger pulse waveform length’ (FPWL), and 'reaction time' (RT). Under variation of the cut-off point for decision, different rates of false-positive (classification of an *irrelevant* trial as *probe*) and false-negative outcomes (classification of a *probe* trial as *irrelevant*) are obtained. This is illustrated by ROC curves. Table 3 shows the areas under the ROC curves and their confidence intervals for each of the single measures and for their optimal-weight combination. Values were calculated separately for both experimental groups.

*** insert table 3 about here ***

For the 'lie' group, EDA_1 and pHR were the best predictors, followed by RLL and EDA_2; all measures contributed to item type discrimination. However, in the 'truth' group, only EDA_1 had a remarkable predictive value; EDA_2 and RT just reached significance, whereas the predictive value of the cardiovascular and respiratory measures missed significance.
Figure 5 shows the ROC curves with the measures EDA_1, EDA_2, RLL, pHR, FPWL, and RT after stepwise inclusion for both experimental groups.

*** insert figure 5 about here ***

Area values obtained from stepwise inclusion were 0.783 for the 'lie' group and 0.637 for the 'truth' group. For a visual comparison, ROC curves with singular inclusion of EDA_1 are included in figure 5. For the 'truth' group, inclusion of EDA_1 almost reached the area under the curve that was obtained from stepwise inclusion. However in the 'lie' group, a considerable amount of incremental validity was obtained from the second electrodermal response, EDA_2, as well as from the respiratory and cardiovascular measures.

Discussion

The present CIT study aimed at separating the psychophysiological responses associated with deceptive behavior from those related to orienting, which are typically coupled in the detection of concealed information. In a laboratory mock-crime experiment with all subjects having 'guilty' knowledge, half of them were instructed to answer truthfully, whereas the others had to conceal their knowledge; subjects' answers were delayed by four seconds after item presentation. Physiological results show that, for subjects concealing their specific knowledge, first and second electrodermal components, respiration line length, and phasic heart rate were largely different between known (*probe*) and unknown (*irrelevant*) objects. For subjects answering truthfully, this difference was found for the electrodermal but not for the respiration and HR measures.

Physiological and behavioral measures as indices of sub-processes of deceptive behavior

Different response patterns were expected for the different measures under variation of both instruction ('truth' vs. 'lie') and trial type (*probe* vs. *irrelevant*).

Several previous studies based on <u>EDA</u> found that delayed-answer paradigms did not improve detection accuracy; in these studies, the electrodermal measures yielded *either no better differentiation (interval following the question) or less differentiation (interval following the answer)* (Furedy et al., 1991). In the present study, the delayed response window led to two distinct electrodermal responses. The first component, which directly followed the item (and question) presentation, differed strongly between *probe* and *irrelevant* trials in the 'truth' and the 'lie' group. This difference is supposed to depend on the orienting response, but as reflected by the Probe by

Group interaction, it also reflects instruction-related processes; whether to deny or to admit scenario-related knowledge. This fact may be explained by an influence of the intention to deceive, including response preparation, or by enhancement of the orienting reflex with a greater subjective relevance of the known item (Furedy and Ben-Shakhar, 1991). However, as a consequence of the answer delay, the deceptive or truthful answering behavior itself is unlikely to be the origin of this differential response.

The second electrodermal component appeared four seconds subsequent to the first and is supposed to reflect mainly answer-related processes. This component also showed a significant *probe*-vs.-*irrelevant* effect in both groups with the greater effect being associated with concealing knowledge; this is in accordance with the findings of Furedy et al. (1988) and Furedy and Ben-Shakhar (1991). A different sensitivity of the two EDA components to the instruction to conceal knowledge was not proven statistically; thus, the EDA results have to be regarded, in sum, as inconclusive with respect to the identification of underlying processes. Greater effect sizes however might possibly infer that the second response component, which has been extracted out of a single skin conductance reaction by the delay in answering, is more dependent on deceptive behavior than the first, yet statistical evidence is missing. If this was the case, then the small, but significant differential response in the 'truth' group might also be due to motivational instructions or other confounds. Subjects were told that they would receive a monetary reward if they 'performed well' in this task, without a further specification; hence, it can be assumed that they focused on identifying the known object within each category among the other objects. If so, giving the "yes" answer would have a greater relevance with respect to the subjects' motivation. Another possible explanation for the differential response in the 'truth' condition might be a generally greater physiological impact of the verbal "yes" answers as compared to "no" answers. An affirmative answer, per se, might lead to a greater physiological response than a negation. Furthermore, the lower frequency of required "yes" answers, confounding the 'Probe' factor with the uniqueness of the "yes" answer, might also have contributed to the observed effect in the 'truth' group. A related phenomenon, the *relative novelty* effect, is also caused by different trial-type frequencies (Ben-Shakhar, 1977). This phenomenon, however, refers to the orienting to presented objects and may therefore have contributed to the differential effect observed in the earlier, but not the later component.

The fact that both EDA components showed habituation, and also the decline of the differential (*probe* vs. *irrelevant*) responses, did not contribute to separating the sub-processes of deception. Orienting theory would predict habituation for the first EDA component due to repeated stimulus presentations; for the observed habituation of the second EDA component however, the repetition of the stimuli is less likely the origin. The slower overall habituation of the second component might underpin this; differential responding however, did not habituate differently between the two components.

For the differentiation between deceptive and truthful denial of knowledge (i.e. *probe* and *irrelevant* trials), RLL has been reported to be a sensitive measure (Timm, 1982a; Gamer et al., 2006). The current results show that the difference in RLL responses to known and unknown objects depends on the instruction to answer deceptively. Because respiratory measures were not found different between *probe* and *irrelevant* trials in the 'truth' group, it may be assumed that RLL is affected by the sub-processes of deceptive behavior in a CIT other than orienting[2]. However, using the ten-seconds scoring method, it cannot be decided whether this measure reflects preparation of deception or deceptive action[3].

The intra-trial course of phasic HR shows a more complex pattern than in most previous lie-detection studies (e.g. Gamer et al., 2006; Gödert et al., 2006). The appearance of several phases can be explained by the fact that the onsets of the induced sub-processes of deceptive behavior were split in time. The motor response with its associated accelerations in HR (e.g. Barry, 1996) and the act of deception, including vocalization, were delayed by about four seconds, whereas orienting processes led to immediate physiological reactions. This presumably led to a different overlap of pHR sub-components than in studies with immediate or no verbal answers. Among the subjects answering truthfully, neither the intra-trial HR pattern nor the corresponding fifteen-seconds average of phasic HR differed significantly between *probe* and *irrelevant* trials. However, among the deceptive subjects, the HR course differed markedly between trial types, and averaged pHR was clearly lower in *probe* than in *irrelevant* trials. Thus, pHR, in contrast to skin conductance but similar to RLL, appears to respond differentially to known vs. unknown objects predominantly when knowledge is concealed[4]. Processes, which might have modified pHR, could be the intention to deceive, preparation of deceptive behavior, or deceptive action. Differential pHR response to *probe* vs. *irrelevant* items in the subjects concealing their knowledge was initiated at trial onsets and lasted throughout trials[2]. These findings support the notion that pHR does not simply reflect orienting, which is in line with earlier findings (e.g. Barry and Maltzman, 1985). Gödert et al. (2006) also found differential responses between trial types in their *innocent-aware group* and concluded that the fact of recognition contributed to phasic changes in HR. The present results did not confirm this finding for the condition involving truthful answers. Here, the decelerative HR response appeared to be uniquely associated with deception, which is a novel finding; among the recorded measures, HR was the most conclusive measure to study the influence of the instruction to conceal knowledge. One possible explanation for this finding, besides masking effects, might be that the orienting-related HR deceleration is essentially influenced by a specific (e.g. crime-related) significance of the object. The finding can also be seen in the light of earlier studies which suggested that the decelerative HR response reflects the relatedness of deception with the evolutionary hiding response (Furedy, 1986).

FPWL showed a response pattern with some similarity to pHR with respect to its sensitivity to trial type and instruction group, but it revealed smaller overall effect size estimates. However, FPWL also differed between known and unknown items in subjects answering truthfully, so that it supports the conclusions from pHR only incompletely.

Lie detection using reaction times has recently been studied by Gronau et al. (2005), Walczyk et al. (2003), Seymour et al. (2000), and Vendemia et al.(2005). In accordance with the findings of these authors, the current study showed longer reaction times to *probe* than to *irrelevant* items when subjects were concealing their knowledge. In subjects answering truthfully, a contrary effect of similar estimated size was found. This differential outcome for RT can be used to identify sub-processes by their incremental contribution to RT (Walczyk et al., 2003). In our study, RT was measured from the prompt to answer to the pressing of the key, which has other implications than in studies without a delay in answering. Supposing that time-consuming processes associated with orienting, evaluation, or answer preparation were completed in our study between the item presentation and the delayed prompt to answer, these processes should not contribute to RT measured after the delayed prompt to answer. The fact that mean RT was shorter than in other CIT studies with immediate answering (e.g. Seymour et al., 2000) is in line with this. Longer RTs in *irrelevant* than in *probe* trials in the 'truth' group signify that truthful 'no' answers require more time-consuming mental processes than truthful 'yes' answers. This might be due to longer comparing processes preceding truthful denying as compared to truthful affirmation of knowledge; the identification of a match might need less time than the decision on a mismatch. Another explanation could be that 'yes' answers need less time than 'no' answers. This was however not controlled in this study. Longer RTs in *probe* than in *irrelevant* trials when subjects concealed their knowledge are in line with study results involving immediate answers, but the implications may be different. If the preparation of the answer was completed within the delay, the fact of deceptively answering with 'no' must contribute to RT for reasons other than stimulus evaluation and answer preparation. Hesitation due to uncertainty, or due to emotional factors (e.g. overcoming fear of being detected or embarrassment when deceiving), might be possible explanations.

The 'detection' perspective

The importance and incremental validity of single psychophysiological and behavioral measures for the detection of concealed knowledge has been studied intensively, and the present results for 'lie' group subjects illustrate this importance. In accordance with previous studies on the detection of concealed knowledge (e.g. Ben-Shakhar and Elaad, 2003; Bradley and Janisse, 1981; and Gamer et al., 2006), EDA, RLL, and pHR differentiated between falsely denied recognition of known objects and correct denial of knowledge about new objects. Here, skin conductance - both components taken together - showed the greatest effect, with the size of the effect matching that obtained in previous studies (Ben-Shakhar and Elaad, 2003), and was followed by pHR and RLL. The relatively high power of pHR may have been achieved by scoring the average change in HR within ten seconds after trial onset, compared to baseline, whereas other authors scored the maximum change in HR within the same time window (Bradley and Janisse, 1981; Gamer, 2006). Also, more speculatively, delaying the answers might have reduced error variance (by changing phases of overlapping influences).

FPWL appeared to be a weak indicator of concealed information. In comparison with electrodermal, respiratory, and ECG measures, only few authors (Podlesny and Raskin, 1978; Podlesny and Kircher, 1999; Hirota et al., 2003) referred to pulse wave in the context of deception and its detection. Although FPWL differed between *probe* and *irrelevant* trials, its contribution to the differentiation between item types was small. Yet, this finding contradicts recent findings by Elaad and Ben-Shakhar (2006), who emphasized the importance of this measure for the detection of deception. Similarly, reaction times contributed little to differentiation of item type.

In research on the detection of concealed knowledge, the psychophysiological and behavioral correlates of orienting, object recognition, and responses in subjects answering truthfully are also of interest. The differentiation of truthful 'yes' vs. truthful 'no' in truthful subjects leads to results which are considerably different from those obtained by the differentiation of deceptive 'no' vs. truthful 'no' in the detection of concealed knowledge. In the group answering truthfully, most of the differentiation of item type that could be achieved by a combination of all physiological measures was yielded by the first electrodermal response component (measured immediately after item presentation). Cardiovascular measures did not, and respiration did only marginally contribute to the differentiation of item type when subjects admitted their deed-related knowledge. Reaction times in the truthful group yielded the same, small amount of incremental information as in the deceptive group, but the effect was in the opposite direction.

Conclusions

The current results show that the delay in answering, although it did not improve detection of concealed information, was informative under a psychophysiological viewpoint. First, two different EDA response components were elicited by the presenation of the items and the answers of the subjects.. Delaying the subjects' answers in combination with applying a response decomposition procedure, made it possible to investigate both electrodermal responses separately. Data from these components however were inconclusive with respect to identifying different mental processes elicited in the CIT.

Second, the present study supports the notion that respiratory and cardiovascular measures (RLL, pHR, and possibly FPWL) are determined by deception-related processes rather than by orienting; they may thus reflect different sub-processes involved in a CIT than EDA. Particularly pHR revealed a conclusive pattern of responses which suggests a relatedness of this measure to deception.

Third, the fact of information concealment, compared with truthful answering, clearly improves differentiation between known and unknown objects. This is especially evident when the differentiation of item type is based not only on the EDA response to the item presentation but rather on a linear combination of electrodermal, cardiovascular, and respiratory measures.

Future research needs to further disentangle intentional and preparatory components from orienting. This might be achieved by the differentiation of truthful vs. deceptive answers in a within-subject design, with an instruction varying between trials. Delayed timing of this instruction with respect to item presentation may help to separate the mentioned components. This might also contribute to elucidating as to why differential responses to *probe* vs. *irrelevant* item presentations are greater when subjects are instructed to conceal their knowledge.

In order to identify other sub-processes such as stimulus classification, the frequency of presentations of known and unknown objects could be modified. This may provide insights into the influences of presentation frequencies on differential responses.

Figure Captions

Figure 1. Presentation of the Concealed Information Test with question and item picture. Fields with question marks besides the depicted item appear with a four-seconds delay after item presentation and prompt the subject to answer. After the key press, a 'yes' or 'no' text (reflecting the subject's answer) replaces the question marks. (Translation of the german question text: 'Has this household object been in the room during your deed?')

Figure 2. Effect sizes for the differential responses to *probe* vs. *irrelevant* items: For the 'truth' and the 'lie' group, Cohen's d is depicted for first electrodermal reaction (EDA_1), second electrodermal reaction (EDA_2), phasic heart rate (pHR), respiration line length (RLL), finger pulse waveform length (FPWL), and reaction times (RT).

Figure 3. Grand means of skin conductance reactions in *scenario-related* trials for the 'truth' group (left) and the 'lie' group (right). Conductance courses are shown for *probe* trials (solid lines) and *irrelevant* trials (dashed lines).

Figure 4. Second-per-second values of heart rate following trial onsets for the 'truth' group (left) and the 'lie' group (right). Phasic heart rate is depicted for *probe* trials (solid lines) and *irrelevant* trials (dashed lines).

Figure 5. ROC curves for 'truth' group and 'lie' group with the predictors EDA_1, EDA_2, pHR, RLL, FPWL, and RT after stepwise inclusion (Wald statistic), and with inclusion of EDA_1 only.

Tables

Table 1

Means and standard errors of means (SEM) of raw scores for each data channel.

Probe vs. *irrelevant* items are contrasted separately for both groups.

		'truth' group				'lie' group			
		probe items		*irrelevant* items		*probe* items		*irrelevant* items	
		Mean	SEM	Mean	SEM	Mean	SEM	Mean	SEM
EDA_1 [nS]		171	27	134	24	259	44	145	25
EDA_2 [nS]		241	63	228	70	325	58	256	49
RLL [arb. units]		2851	404	2890	405	2823	312	3169	351
pHR [1/min]		1.07	0.29	1.33	0.15	-1.03	0.31	1.24	0.20
FPWL [arb. units]		158	13	164	14	158	23	169	24
RT [ms]		598	30	649	33	903	56	843	52

Table 2

Physiological measures included in the binary logistic regression following the Wald statistic: components of electrodermal activity (EDA_1, EDA_2), phasic heart rate (pHR), respiration line length (RLL), reaction time (RT), and finger pulse waveform length (FPWL). Optimal weights (β), standard errors (SE) of β, Wald statistic and significance level (*p*) are listed for 'lie' group and 'truth' group.

measure	**β**	**SE**	**Wald**	***p***
'lie' group				
EDA_1	0.552	0.062	79.71	< 0.001
EDA_2	0.404	0.061	43.15	< 0.001
pHR	-0.476	0.062	58.58	< 0.001
RLL	-0.725	0.072	102.68	< 0.001
RT	0.324	0.061	28.38	< 0.001
FPWL	-0.165	0.061	7.34	< 0.01
constant	-1.368	0.064	456.71	< 0.001
'truth' group				
EDA_1	0.380	0.055	47.40	< 0.001
EDA_2	0.180	0.059	9.37	< 0.01
RLL	-0.137	0.060	5.27	< 0.05
RT	-0.269	0.064	17.58	< 0.001
constant	-1.148	0.057	406.69	< 0.001

Table 3

Area under the receiver operating characteristic (ROC) curves and 95% confidence intervals for differentiation between *probe* and *irrelevant* items on a single-trial basis. Values are listed for inclusion of each single measure and for stepwise inclusion (Wald statistic). Results for 'lie' group and 'truth' group are presented in separate columns.

included parameters	**area under the ROC curve and 95% confidence intervals**			
	'lie' group		**'truth' group**	
	area	confidence interval	area	confidence interval
single measures:				
EDA_1	0.669	0.641 - 0.697	0.612	0.582 - 0.642
EDA_2	0.617	0.588 - 0.646	0.548	0.517 - 0.579
pHR	0.669	0.641 - 0.696	0.510	0.479 - 0.541
RLL	0.641	0.613 - 0.669	0.514	0.484 - 0.545
RT	0.570	0.541 - 0.599	0.565	0.534 - 0.595
FPWL	0.571	0.542 - 0.600	0.516	0.486 - 0.546
after stepwise inclusion:				
	0.783	0.760 - 0.807	0.637	0.608 - 0.665

Footnotes

[1] Group main effects could also be investigated on the basis of raw values. However, a mean difference between the two instructional groups was not of interest.

[2] The finding that RLL was not sensitive to the Probe factor for the 'truth' group might be relativized by a replication study using a larger sample; this would however not contradict the finding of a different responsivity between the 'truth' and the 'lie' group.

[3] Analyses of RLL and pHR did not allow a decomposition of responses or a response scoring with respect to the different time windows. Decomposition is helpful in the analysis of EDA, but for respiratory and cardial measures no response templates are known, which would allow similar decomposition procedures. For pHR, an additional second-per-second analysis revealed significant differences between *probe* and *irrelevant* trials from the third to the fifteenth second after trial onset.

[4] HR deceleration to *novel* stimuli could also be expected in the 'truth' condition. In contrast to previous studies (e.g. Gödert et al., 2001), this was not observed in the present study. Possibly, such a deceleration (which should follow object presentation), was masked by the four-seconds delayed motor response component (vocalization and key press) in this study.

References

Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of Mathematical Psychology*, *12*, 387-415.

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

Barry, R. J., Maltzman, I. (1985). Heart rate deceleration is not an orienting reflex; heart rate acceleration is not a defensive reflex*. Pavlovian Journal of Biological Science*, *20*, 15-28.

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

Ben-Shakhar, G. (1985). Standardization within individuals: a simple method to neutralize individual differences in skin conductance. *Psychophysiology*, *22*, 292-299.

Ben-Shakhar, G., Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, *88*, 131-151.

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

Bradley, M. T., Rettinger, J. (1992). Awareness of crime-relevant information and the Guilty Knowledge Test. *Journal of Applied Psychology, 77*, 55-59.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. SanDiego, CA: McGraw-Hill.

Dawson, M. E. (1980). Physiological detection of deception: Measurement of responses to questions and answers during countermeasure maneuvers. *Psychophysiology*, *17*, 8-17.

Elaad, E., Ben-Shakhar, G. (2006). Finger pulse waveform length in the detection of concealed information. *International Journal of Psychophysiology*, *61*, 226-234.

Furedy, J. J. (1986). Lie detection as psychophysiological differentiation: Some fine lines. In M.Coles, E. Donchin, A S. Porges (Eds.), Psychophysiology: Systems, Processes, and Applications - A Handbook. Guilford.

Furedy, J. J., Ben-Shakhar, G. (1991). The Roles of Deception, Intention to Deceive, and Motivation to Avoid Detection in the Psychophysiological Detection of Guilty Knowledge. *Psychophysiology*, *28*, 163-171.

Furedy, J. J., Davis, C., Gurevich, M. (1988). Differentiation of Deception as a Psychological Process: A Psychophysiological Approach. *Psychophysiology*, *25*, 683-688.

Furedy, J. J., Gigliotti, F., Ben-Shakhar, G. (1994). Electrodermal differentiation of deception: The effect of choice versus no choice of deceptive items. *International Journal of Psychophysiology*, *18*, 13-22.

Furedy, J. J., Posner, R. T., Vincent, A. (1991). Electrodermal differentiation of deception: Perceived accuracy and perceived memorial content manipulations. *International Journal of Psychophysiology*, *11*, 91-97.

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*, 76-87.

Gödert, H. W., Gamer, M., Keth, A., Rill, H. G., Vossel, G. (2006). Phasic Heart Rate in the Guilty Actions Test. The effects of crime-relevant knowledge and deception on phasic heart rate in the Guilty Actions Test. Manuscript (submitted).

Gödert, H. W., Rill, H. G., Vossel, G. (2001). Psychophysiological differentiation of deception: The effects of electrodermal lability and mode of responding on skin conductance and heart rate. *International Journal of Psychophysiology*, *40*, 61-75.

Gronau, N., Ben-Shakhar, G., Cohen, A. (2005). Behavioral and Physiological Measures in the Detection of Concealed Information. *Journal of Applied Psychology*, *90*, 147-158.

Hirota, A., Sawada, Y., Tanaka, G., Nagano, Y., Matsuda, I., Takasawa, N. (2003). A new index for *psychophysiological* detection of deception: applicability of normalized pulse volume. *Japanese Journal of Physiological Psychology and Psychophysiology*, *21*, 217-230.

Horneman, C.J., O'Gorman, J.G. (1985). Detectability in the Card Test as a function of the subject's verbal response. *Psychophysiology, 22*, 330-333.

Kircher, J. C., Raskin, D. C. (2003). *The computerized polygraph system II (Software Version 4.01)*. Scientific Assessment Technologies, Inc., 2532 Chadwick Street, Salt Lake City, Utah, USA 84106.

Kugelmass, S., Lieblich, I., Bergman, Z. (1967). The role of "lying" in psychophysiological detection. *Psychophysiology, 3*, 312-315.

Lim, C. L. (1999). Dynamics of SCR, EEG, and ERP activity in an oddball paradigm with short interstimulus intervals. *Psychophysiology*, *36*, 543-551.

Lim, C.L., Rennie, C., Barry, R., Bahramali, H., Lazzaro, I., Manor, B., Gordon, E. (1997). Decomposing skin conductance into tonic and phasic components. *International Journal of Psychophysiology*, *25*, 97-109.

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

Lykken, D. T., Venables, P. H. (1971). Direct Measurement of Skin Conductance: A Proposal for Standardization. *Psychophysiology*, *8*, 656-672.

Meijer, E. H., Smulders, F. T. Y., Johnston, J. E., Merckelbach, H. L. G. J. (2007). Combining skin conductance and forced choice in the detection of concealed information. *Psychophysiology, 44*, 814-822.

Podlesny, J. A., Kircher, J. C. (1999). The Finapres (Volume Clamp) Recording Method in Psychophysiological Detection of Deception Examinations. Experimental Comparison With the Cardiograph Method. *Forensic Science Communications, 1* (3). Retrieved November, 02, 2006, from http://www.fbi.gov/hq/lab/fsc/backissu/oct1999/podlsny1.htm

Podlesny, J. A., Raskin, D. C. (1978). Effectiveness of Techniques and Physiological Measures in the Detection of Deception. *Psychophysiology*, *15*, 344-359.

Rosnow, R. L., Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Psychological Methods*, *1*, 331-340.

Seymour, T. L., Seifert, C. M., Shafto, M. G., Mosmann, A. L. (2000). Using response time measures to assess "guilty knowledge". *Journal of Applied Psychology*, *85*, 30-37.

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

Timm, H. W., (1982b). Effect of altered outcome expectancies stemming from placebo and feedback treatments on the validity of the guilty knowledge technique. *Journal of Applied Psychology*, *67*, 391-400.

Velden, M., Wölk, C. (1987). Depicting cardiac activity over real time: A proposal for standardization. *Journal of Psychophysiology*, *1*, 173-175.

Vendemia, J. M. C., Buzan, R. F., Simon-Dack, S. L. (2005). Reaction time of motor responses in two-stimulus paradigms involving deception and congruity with varying levels of difficulty. *Behavioural Neurology*, *16*, 25-36.

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*, 461-466.

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*, 309-318.

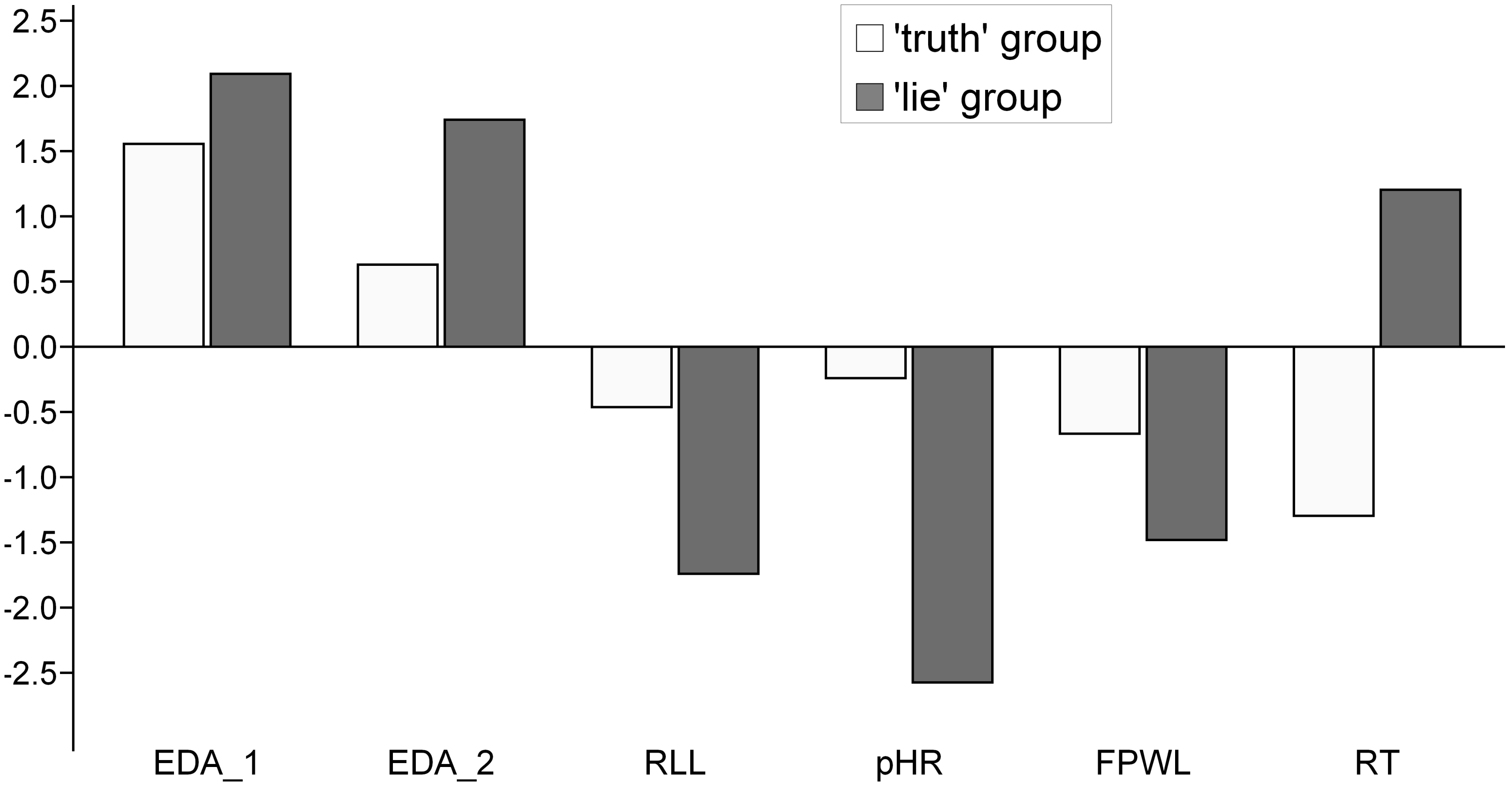
Vincent, A., Furedy, J. J. (1992). Electrodermal differentiation of deception: Potentially confounding and influencing factors. *International Journal of Psychophysiology*, *13*, 129-136.

Walczyk, J. J., Roper, K. S., Seemann, E., Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: response time as a cue to deception. *Applied Cognitive Psychology*, *17*, 755.

Wölk, C., Velden, M., Zimmermann, U., Krug, S. (1989). The interrelation between phasic blood pressure and heart rate changes in the context of the 'baroreceptor hypothesis.'. *Journal of Psychophysiology*, *3*, 397-402.

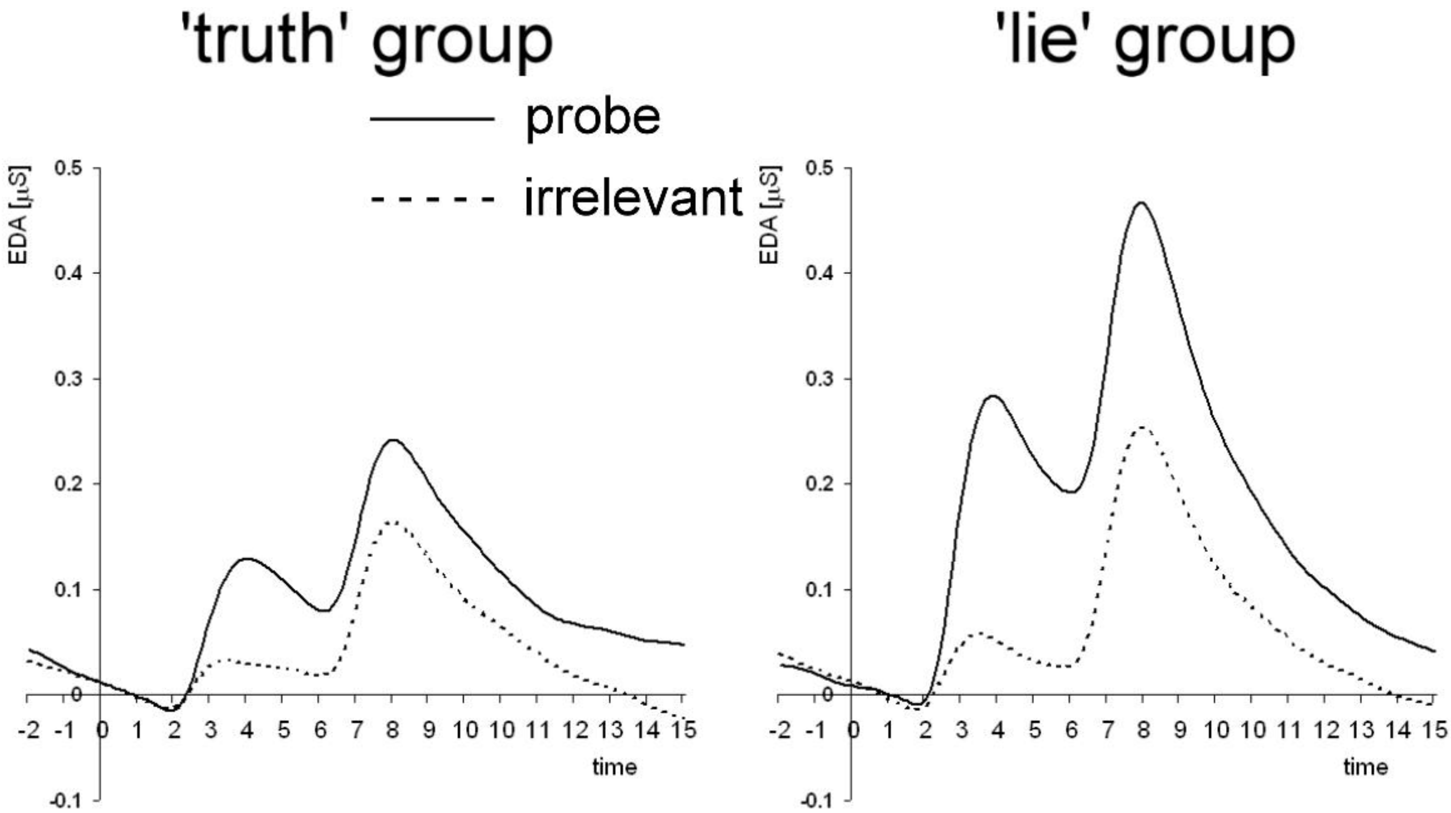
figure1.tif

figure2.tif

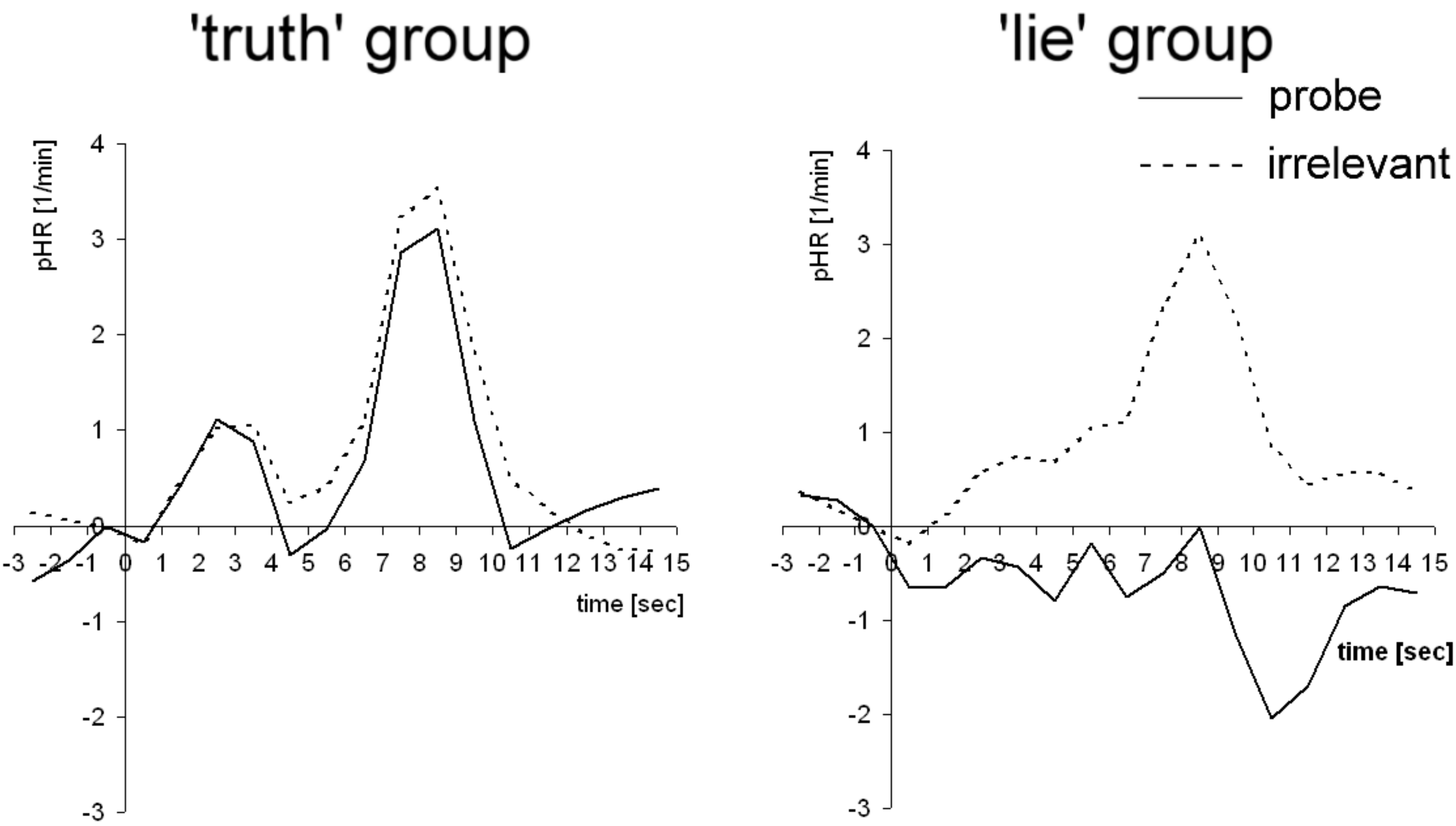


2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
'truth' group
'lie' group
EDA_1
EDA_2
RLL
pHR
FPWL
RT

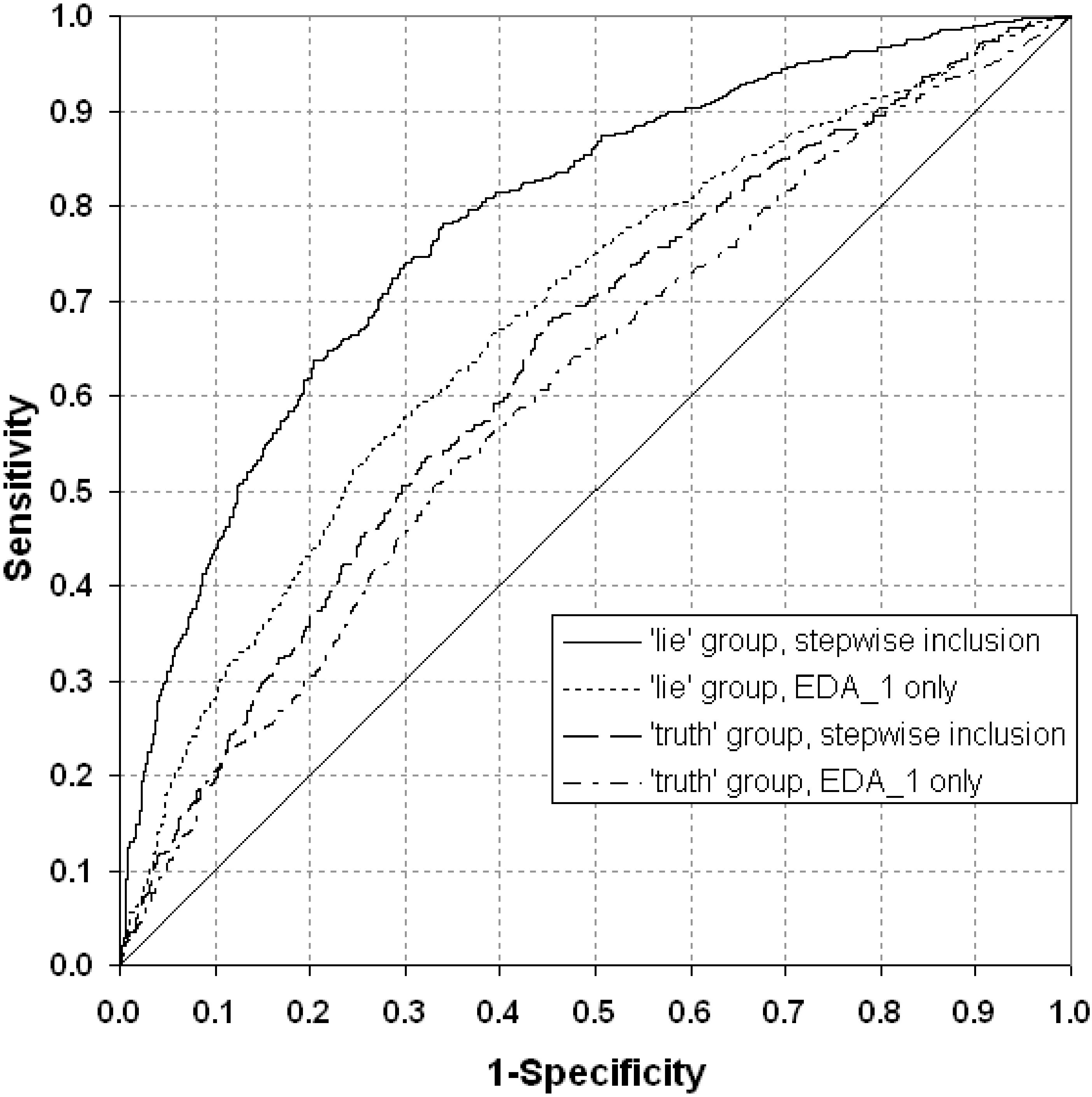
figure3.tif



'truth' group
'lie' group
probe
irrelevant
EDA [µS]
0.5
0.4
0.3
0.2
0.1
0
-0.1
-2 -1 0 1 2 3 4 5 6 7 8 9 10 10 11 12 13 14 15
time



'truth' group
'lie' group
probe
irrelevant
pHR [1/min]
time [sec]
pHR [1/min]
time [sec]



1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Sensitivity
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1-Specificity
'lie' group, stepwise inclusion
'lie' group, EDA_1 only
'truth' group, stepwise inclusion
'truth' group, EDA_1 only