Using pretest data to screen low-reactivity individuals in the autonomic-based concealed information test

Matsuda, I.; Ogawa, T.; Tsuneoka, M.; Verschuere, B.

DOI
10.1111/psyp.12328

Publication date
2015

Document Version
Final published version

Published in
Psychophysiology

Citation for published version (APA):
Matsuda, I., Ogawa, T., Tsuneoka, M., & Verschuere, B. (2015). Using pretest data to screen low-reactivity individuals in the autonomic-based concealed information test. *Psychophysiology*, *52*(3), 436-439. https://doi.org/10.1111/psyp.12328

Psychophysiology, 52 (2015), 436–439. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/psyp.12328

PSYCHOPHYSIOLOGY

BRIEF REPORT

Using pretest data to screen low-reactivity individuals in the autonomic-based concealed information test

IZUMI MATSUDA,[a] TOKIHIRO OGAWA,[a] MICHIKO TSUNEOKA,[a] AND BRUNO VERSCHUERE[b,c,d]

[a]National Research Institute of Police Science, Chiba, Japan
[b]Department of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands
[c]Department of Experimental, Clinical, & Health Psychology, Ghent University, Ghent, Belgium
[d]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

Abstract

The concealed information test (CIT) can be used to assess whether an individual possesses crime-related information. However, its discrimination performance has room for improvement. We examined whether screening out participants who do not respond distinctively on a pretest improves the diagnosticity of a mock-crime CIT. Before conducting the CIT, we gave a pretest to 152 participants, 80 of whom were assigned as guilty. Pretest screening significantly improved the diagnostic value of the mock-crime CIT; however, it also led to a substantial number of undiagnosed participants (33.6%). Pretest screening holds promise, but its application would benefit from dedicated measures for screening out participants.

Descriptors: Concealed information test (CIT), Screening, Autonomic responses, Memory detection, Deception, Lie detection

The concealed information test (CIT), also known as the guilty knowledge test (GKT), typically uses physiological responses to estimate whether an individual possesses crime-related information (Lykken, 1959; for a review, see Verschuere, Ben-Shakhar, & Meijer, 2011). In a CIT, one crime-relevant item is presented among a series of irrelevant items. Those items are selected so that innocent people will be unable to distinguish the crime-relevant item from the irrelevant ones. If the physiological responses to the relevant item differ from those to the irrelevant ones, the examiner may infer that the individual possesses knowledge about the relevant item. In Japan, the CIT has been officially and widely used in criminal investigations (Matsuda, Nittono, & Allen, 2012).

However, the discriminatory performance of the CIT is not perfect. Some individuals do not react differently to a crime-relevant item even if they recognize it (Matsuda et al., 2012). To improve discrimination performance, previous studies have tried to combine different types of measures, for example, using several autonomic measures (Gamer, Verschuere, Crombez, & Vossel, 2008), or amalgamating autonomic measures with event-related potentials (ERPs; Matsuda, Nittono, & Ogawa, 2011).

A recent study using a CIT based on reaction times (Noordraven & Verschuere, 2013) indicated that discrimination performance can also be improved by screening out participants who may not react to the CIT. The participants were asked to press the “yes” key to target items and “no” key to all other items (including the crime-relevant items embedded among several irrelevant items). This was based on the reasoning that knowledgeable participants respond similarly to target and crime-relevant items; that is, both lead to increased reaction times compared with the irrelevant items. Thus, participants were eliminated when they failed to show a marked target response. This screening improved the diagnostic efficiency of the crime CIT (i.e., the area under a receiver operating characteristic curve (AUC), from 0.87 to 0.95, with 21.4% of the participants eliminated). However, this method does not allow for independent screening before conducting the crime CIT.

The present study explores a new screening procedure for the more widely researched and forensically applied autonomic-based CIT. We examined the results of a screening based on a participant’s reactivity in a pretest, which is usually performed before the crime CIT in forensic applications (Matsuda et al., 2012). Although the main purpose of the pretest is to demonstrate subsequent crime CIT, the result may also be used to capture participants’ reactivity prior to conducting the actual crime CIT. The purpose of the present study was to examine the effectiveness of screening participants based on their pretest autonomic responses. We predicted that this screening would improve the discrimination performance in the subsequent crime CIT.

In a supplementary analysis, we tested another means of using pretest data—that proposed by Matsuda, Hirota, Ogawa, Takasawa, and Shigemasu (2006). These authors proposed employing weighted crime CIT responses according to pretest response tendencies and found that such weighting significantly improved discrimination performance.

Address correspondence to: Izumi Matsuda, Ph.D., National Research Institute of Police Science, 6-3-1, Kashiwanoha, Kashiwa, Chiba, Japan. E-mail: izumi@nrips.go.jp

Method

Participants

We ran secondary analyses on the data from participants in the Ogawa, Matsuda, and Tsuneoka (2013) study. There was a total of 152 participants (81 men and 71 women, 18–59 years old, $M = 36.6$, $SD = 11.0$), who all worked in police organizations. Of these, 80 and 72 participants were randomly assigned to the guilty and innocent groups, respectively. All participants gave their informed consent prior to the experiment.

Examiners

Thirty-six Japanese polygraphers conducted the tests. All the polygraphers belonged to either the Forensic Science Laboratories or the National Research Institute of Police Science in Japan.

Stimuli

Five numbers (3, 4, 5, 6, and 7) were used in the pretest, and five accessories (necklace, earrings, ring, watch, and brooch) were employed in the mock-crime CIT. Each item was presented both vocally by the examiner and visually on screen for 3 s as a string of letters.

Procedure

Each participant was left alone with an envelope including role instructions (guilty or innocent). Guilty participants stole a ring (i.e., a relevant item). Innocent participants merely stayed in the room and waited.

The participants then took a pretest. They selected one of five cards and memorized the number on it (i.e., a relevant item). They adopted a sitting position such that they would have no eye contact with the experimenter. Physiological measurement devices were attached, and then the pretest began. For every pretest block, each of the five numbers was presented once in a random order with an interstimulus interval of 25–40 s. The participants were asked to reply "no" to all the items. The block of five numbers was repeated five times, with the numbers in each block presented in a different order to eliminate serial-order effects.[1]

Subsequently, the participants took a crime CIT, in which the name of each of the five accessories was presented. Apart from the item content, the crime CIT procedure was identical to that of the pretest.

Recording and Reduction

The following four autonomic responses were recorded using a polygraph system (PTH-347, TEAC, Japan). The sampling rate was 1000 Hz.

1. Respiration. A carbon film-type transducer was placed around the participant's abdomen. The output waveforms were resampled at 20 Hz with a low-pass filter of 4 Hz. Then, the respiration line length was calculated by summing the change in values for each time unit for 10 s after the participant's reply.
2. Skin conductance. Two Ag/AgCl disposable electrodes (PPS-EDA, TEAC) covered with 0.05 M NaCl electrolyte were placed on the volar side of the distal phalanges of the index and middle fingers of the nondominant hand. These electrodes were connected to a skin conductance unit, which supplied a constant voltage of 0.5 V. Skin conductance response was identified as the maximum-minimum difference in the wave occurring 0.5–5 s after the stimulus onset.
3. Heart rate. An electrocardiogram was recorded with a standard lead II configuration. The signal was amplified with a time constant of 0.1 s and a low-pass filter of 30 Hz. R-R intervals were converted into heart rate with real-time scaling. The heart rate was averaged over the 6- to 16-s poststimulus period.
4. Normalized pulse volume. A near-infrared light-emitting diode and a phototransistor were placed at opposite sides of the fourth finger of the nondominant hand. The phototransistor output was amplified with a time constant of 0.3 s and without filters to generate the normalized pulse volume (Sawada, Tanaka, & Yamakoshi, 2001), which was averaged over the 6- to 16-s poststimulus period.

Effect Size

For each participant, measure, and test type, an effect size (Cohen's *d*) was calculated between the relevant item and four irrelevant items so as to evaluate each test participant's reactivity. The signs of the *d*s were reversed for respiration line length, heart rate, and normalized pulse volume; this was because those measures typically decrease for a relevant item compared with irrelevant items when participants recognize the relevant item. The *d*s of the four measures were then averaged; hereafter, we refer to this as the "combined *d*" (cf. *z* score averaging reviewed in Matsuda et al., 2012). The 95% bootstrap confidence intervals of the *d*s were calculated using the MATLAB bootci function.

Analytic Plan

First, to evaluate whether reactivity in the pretest could predict the results in the crime CIT, we calculated effect-size correlations between the pretest and crime CIT. These correlations were determined separately for the guilty and innocent groups.

Second, we screened out participants whose combined *d* in the pretest was less than an a priori defined cutoff. From the four cutoffs examined by Noordraven & Verschuere (2013)—$d < 0$, 0.2, 0.5, and 0.8, we selected the $d < 0.2$ and 0.5 cutoffs, because they provide optimal individual diagnostics when equally balancing sensitivity and specificity or when valuing specificity more highly than sensitivity (see Table 2 of Noordraven and Verschuere, 2013), respectively. For each a priori defined cutoff, we evaluated the crime CIT discrimination performance by using an AUC calculated with the MATLAB perfcurve function. We also calculated the 95% bootstrap confidence interval of the AUC differences before and after screening with each cutoff, following Rosenfeld's (2011) method of evaluating ERP differences.

Supplementary Analysis

We also individually weighted the physiological measures in the crime CIT according to pretest response tendencies.[2] First, we calculated the pretest effect size r ($-1 \leq r \leq 1$) for each participant with each measure. Then, we weighted the *d* of the

1. Usually, CIT studies present a buffer item at the beginning of each block to avoid the primacy effect. Instead of using buffer items, we countered the primacy effect by means of the described protocol.

2. We improve the weighting method of Matsuda et al. (2006) in order to apply it to the effect sizes.

Table 1. *Descriptive Statistics of Cohen's ds and AUCs*

	Mean	95% confidence interval	Correlation with *d*s in pretests
*d*s in pretests (guilty group, $N = 80$)			
Respiration line length	0.680	0.519–0.865	
Skin conductance response	0.884	0.741–1.02	
Heart rate	0.736	0.598–0.878	
Normalized pulse volume	0.465	0.361–0.585	
Combined	0.691	0.622–0.769	
*d*s in pretests (innocent group, $N = 72$)			
Respiration line length	0.515	0.381–0.654	
Skin conductance response	0.882	0.686–1.10	
Heart rate	0.500	0.371–0.637	
Normalized pulse volume	0.490	0.356–0.634	
Combined	0.597	0.522–0.681	
*d*s in mock-crime CITs (guilty group, $N = 80$)			
Respiration line length	0.582	0.405–0.771	0.470*
Skin conductance response	0.893	0.676–1.19	0.296*
Heart rate	0.878	0.707–1.06	0.471*
Normalized pulse volume	0.402	0.296–0.515	0.248*
Combined	0.693	0.594–0.803	0.386*
*d*s in mock-crime CITs (innocent group, $N = 72$)			
Respiration line length	0.033	–0.069–0.122	–0.104
Skin conductance response	–0.088	–0.197–0.035	–0.079
Heart rate	–0.036	–0.159–0.064	–0.137
Normalized pulse volume	0.024	–0.069–0.120	0.067
Combined	–0.015	–0.073–0.044	–0.015
AUCs of mock-crime CITs (no screening, $N = 152$)			
Respiration line length	0.747	0.650–0.819	
Skin conductance response	0.826	0.747–0.896	
Heart rate	0.848	0.774–0.899	
Normalized pulse volume	0.722	0.632–0.795	
Combined	0.919	0.857–0.953	
AUCs of mock-crime CITs (screening using $d < 0.5$ cutoff, leaving $N = 101$)			
Respiration line length	0.825	0.706–0.910	
Skin conductance response	0.882	0.801–0.937	
Heart rate	0.868	0.772–0.927	
Normalized pulse volume	0.740	0.641–0.846	
Combined	0.967	0.916–0.989	

*$p < .05$.

measure m for the participant i in the crime CIT with the following value calculated from rs in the pretest: $(r_{im} - (-1))/\{(r_{i1} - (-1)) + (r_{i2} - (-1)) + (r_{i3} - (-1)) + (r_{i4} - (-1))\}$. Finally, we summed the weighted *d*s across measures; these were used for calculating the AUC.

Results

Effect Size and Correlation

Table 1 shows the means[3] and correlations of Cohen's *d*s. For the guilty participants, pretest reactivity was significantly related to crime CIT reactivity for all measures. No such relationship was found for the innocent group.

Screening

Figure 1 shows the AUCs and the proportions of removed participants with the cutoffs of the combined $d < 0$–1 in the pretest. The 95% confidence intervals of the AUC differences before and after screening with a priori defined $d < 0.2$ and 0.5 cutoffs were -0.059–0.071 (13.2% of participants were eliminated) and 0.001–0.105 (33.6% were eliminated). Table 1 presents the AUCs before and after screening with the $d < 0.5$ cutoff.

Supplementary Analysis

The AUC of the weighting method was 0.920 (95% confidence interval: 0.868–0.958). This method did not improve the performance of the crime CIT.

Discussion

The present study is the first to examine whether the diagnosticity of a crime CIT can be improved by screening out individuals who fail to show distinctive responses in a pretest. We found that pretest reactivity was able to predict mock-crime CIT reactivity, which admits the possibilities for effective screening. Individuals who did not show a substantial pretest effect were screened out of the crime CIT. Compared with before screening, the crime CIT discrimination showed significant improvement using the $d < 0.5$ cutoff. Thus, the pretest data were effective in detecting low-reactivity participants; however, they caused many participants to be excluded from the discrimination. Such individuals would produce

3. Table 1 indicates that the pretest reactivity was larger for the guilty group than for the innocent group. However, a t test of the combined pretest *d*s showed only marginally significant differences: $t(150) = 1.88$, $p = .062$, $d = 0.308$.

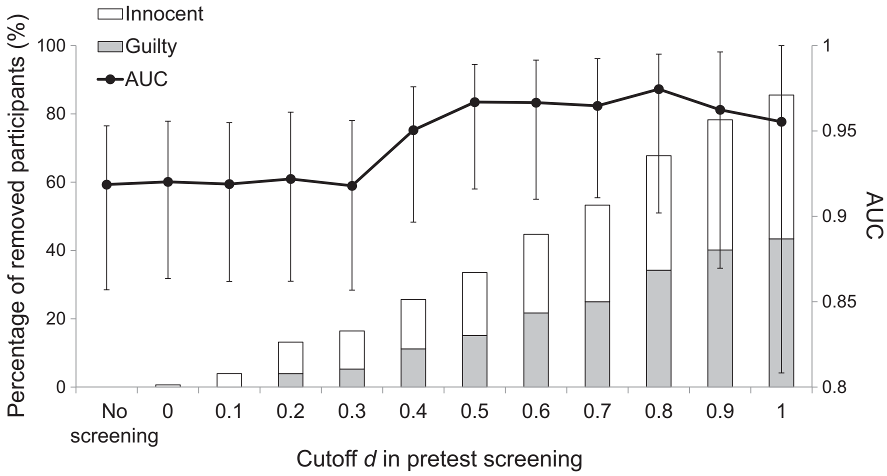

Figure 1. The line graph shows AUCs and their 95% confidence intervals in the mock-crime CIT by screening participants using combined $d < 0$–1 cutoffs in the pretest. The bar graph shows the percentage of removed participants at each cutoff. The a priori defined cutoffs are highlighted in bold.

inconclusive test results. Another, less drastic, way to use the pretest result would be in interpreting the crime CIT outcome: if a participant showed distinctive responses in the pretest, but not in the crime CIT, an examiner could interpret this to mean that the participant lacked knowledge of the crime-relevant item—rather than supposing the participant to be a low responder.

Before applying the proposed screening method to actual situations, it is necessary to verify the present findings using a field dataset. In field settings, guilty participants would be more motivated to conceal their knowledge than in an experiment, and such a difference in motivation could affect autonomic responses (Ben-Shakhar & Elaad, 2003). More research is needed to identify appropriate field-setting effect-size cutoffs.

The supplementary analysis showed that—contrary to expectations from Matsuda et al. (2006)—weighting each measure according to pretest responses did not work well in the present dataset. The correlation between the pretest and mock-crime CIT here ($r = .386$; $n = 80$) was substantially smaller than that found by Matsuda et al. ($r = .673$; $n = 19$). It would appear that more research is needed to validate the conditions under which the weighting method can effectively improve the diagnostic efficiency of the CIT.

The present study points to the potential value of pretest screening. One way to use the results of the pretest is to refrain from applying the crime CIT to low-reactive individuals—as we did in the present study. One could also use the pretest information to interpret the crime CIT outcome. Finally, future research needs to show whether pretesting can be used to select a CIT (e.g., based on autonomic responses, ERPs, or reaction times) that is more appropriate for each individual.

References

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.

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

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

Matsuda, I., Hirota, A., Ogawa, T., Takasawa, N., & Shigemasu, K. (2006). A new discrimination method for the Concealed Information Test using pretest data and within-individual comparisons. *Biological Psychology*, *73*, 157–164.

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*, 532.

Matsuda, I., Nittono, H., & Ogawa, T. (2011). Event-related potentials increase the discrimination performance of the autonomic-based Concealed Information Test. *Psychophysiology*, *48*, 1701–1710.

Noordraven, E., & Verschuere, B. (2013). Predicting the sensitivity of the reaction time-based Concealed Information Test. *Applied Cognitive Psychology*, *27*, 328–335.

Ogawa, T., Matsuda, I., & Tsuneoka, M. (2013). Accuracy of Concealed Information Test as a memory detection technique: A laboratory study [in Japanese]. *Japanese Journal of Forensic Science and Technology*, *18*, 35–44.

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, UK: Cambridge University Press.

Sawada, Y., Tanaka, G., & Yamakoshi, K. (2001). Normalized pulse volume (NPV) derived photo-plethysmographically as a more valid measure of the finger vascular tone. *International Journal of Psychophysiology*, *41*, 1–10.

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

(Received February 26, 2014; Accepted August 8, 2014)