International Journal of Psychophysiology 81 (2011) 60–63

Contents lists available at ScienceDirect

International Journal of Psychophysiology

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

Short Communication

A novel countermeasure against the reaction time index of countermeasure use in the P300-based complex trial protocol for detection of concealed information ☆

Alexander Sokolovsky *, Jessica Rothenberg, Elena Labkovsky, John Meixner, J. Peter Rosenfeld

Department of Psychology, Northwestern University, Evanston, IL, USA

ARTICLE INFO

Article history:
Received 18 June 2010
Received in revised form 9 March 2011
Accepted 11 March 2011
Available online 31 March 2011

Keywords:
ERP
P300
CTP
Deception detection
Countermeasure
Concealed information

ABSTRACT

A P300 deception detection protocol was tested using simultaneous versus serial countermeasures and stimulus acknowledgment responses. Previously, P300 showed recognition and elevated reaction time identified countermeasure use. Probe-irrelevant P300 differences were significant in both countermeasure groups and control group. Detection rates were 11/12 for controls, 10/12 for serial countermeasure users, and 11/13 for simultaneous countermeasure users. Reaction time detected countermeasure use in serial responders, but not simultaneous responders. The simultaneous response reaction times were indistinguishable from controls.

For almost a century, and with renewed intensity since September 11, 2001, there have been enormous efforts expended by governments and universities to develop an accurate deception test based on sound scientific principles. Both polygraph protocols using the measurements of autonomic nervous system activity (the Comparison Question Test (CQT) and the Concealed Information Test (CIT)) have been alternatively advocated and criticized, as recently summarized in a long report by the National Research Council of the National Academy of Sciences (National Academies Press, 2003). Among the problems with polygraphy raised by the National Research Council report is its potential susceptibility to *countermeasures*. As stated by Honts et al. (1996, p. 84), "Countermeasures are anything that an individual might do in an effort to defeat or distort a polygraph test." The National Research Council report went on to state that "Countermeasures pose a serious threat to the performance of polygraph testing because all the physiological indicators measured by the polygraph can be altered by conscious efforts through cognitive or physical means" (National Academies Press, 2003). More specifically, countermeasures (CMs) are effective against both the polygraphic CQT, (Honts, et al., 2001) as well as against the polygraphic CIT (Ben-Shakhar and Dolev, 1996; Elaad and Ben-Shakhar, 1997; Honts, et al., 1996).

It was hoped and indeed expected that when the P300 ERP component was introduced as the dependent index of recognition in a CIT (Farwell and Donchin, 1991; Rosenfeld, et al., 1991) the CM issue would be resolved. The eminent inventor of the CIT, Lykken (1998, p. 293) put it this way: "Because such potentials are derived from brain signals that occur only a few hundred milliseconds after the GKT alternatives are presented…it is unlikely that countermeasures could be used successfully to defeat a GKT derived from the recording of cerebral signals." (Ben-Shakhar and Elad, 2002 expressed a similar view.) Unfortunately, Rosenfeld et al. (2004) and Mertens and Allen (2008) showed that the original form of the P300-based CIT was vulnerable to CMs, prompting development of a novel P300-based protocol which has thus far resisted previously effective CMs (Mertens and Allen, 2008) in three new studies (Rosenfeld and Labkovsky, 2010; Rosenfeld, et al., 2008; Winograd and Rosenfeld, 2011).

Indeed the novel *complex trial protocol (CTP)* has so far been the *only* physiologically based deception testing protocol reported that is resistant to CMs, and additionally, provides a simple index – reaction time (RT) – of the *use* of a CM by subjects. Thus, the test has so far reliably identified recognition of concealed information as well as the attempt by guilty subjects in a forensic situation to counter the protocol — which likely constitutes additional evidence of a subject's criminal complicity. Moreover, even in the rare cases, occasionally encountered, in which a subject who is instructed to beat the test succeeds in not showing the enhanced P300 indicator of guilty knowledge recognition, his RT index may still give away his attempt at non-cooperation — useful information for enforcement officials.

Therefore, it was most disappointing insight for us to appreciate, based on some pilot data with new subjects, the major threat posed by a new CM that neither we nor any reviewers of our recent studies previously suggested. The new CM, the subject of the present report,

☆ This research was supported by the Defense Academy for Credibility Assessment grants DODP198-P-0001, DODPI04-P-0002, and W74V8h-6-01-0001 awarded to J. Peter Rosenfeld.

* Correspondence author. Tel.: +1 913 653 2741.
E-mail address: alexandersokolovsky2007@u.northwestern.edu (A. Sokolovsky).

0167-8760/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpsycho.2011.03.008

self-evidently threatened the RT index of CM use, but possibly also threatened the more basic P300 index of information recognition afforded by our previously accurate and CM-resistant new P300 protocol. Understanding this issue requires some background.

In the CTP, subjects are presented in each trial with two sequential tasks. The first (and critical) one involves responding immediately to either a probe (crime-relevant or key) or irrelevant stimulus with a single button press (called the “I saw it” response). The meaningful and rare probe but not the irrelevant stimulus elicits the P300 sign of recognition. Thus, the key indicator of guilty knowledge recognition is a large probe P300 as compared to the irrelevant P300. RT to the “I saw it” response *to irrelevant stimuli being countered* will typically be elevated in subjects taught to counter some or all irrelevant stimuli since subjects need to take some moments to recall which CM to do (possibly also after recalling which, if not all, irrelevants need countering). This had to have been the case in our previous CTP studies in which we explicitly instructed subjects to recall and execute the CM *prior to* doing the “I saw it” button response. We so instructed them based on our empirically confirmed belief that if they did the CM *after* the button response, the intended effect of the CM – to make the irrelevant stimulus task-relevant and thus meaningful and fully capable of eliciting a P300 as large as the probe P300 – would be lost, rendering the CM ineffective. The new countermeasure that threatens the CTP utility will be tested in this experiment; it simply involves executing the CM simultaneously with doing the “I saw it” button response.

Thirty-eight individuals were recruited from Northwestern University. All subjects gave written consent to participate and were given $10/h for participation. All subjects had normal or corrected to normal vision. Subjects with contact lenses were asked to wear glasses to the experiment. This experiment was approved by the Northwestern University Institutional Review Board.

Each trial began with a 100 ms baseline period during which pre-stimulus EEG was recorded. A date was then presented in white text on a black background in the middle of the screen for 300 ms in this format: “Dec 17”. Probe dates were birth dates; others were irrelevant. Upon seeing the stimulus, subjects were instructed to press the left button on a two button box using their right hand. Because the button press represented the participant seeing the stimulus whether or not he/she recognized it, this was called the “I saw it” response. Subjects were divided into three groups: *simple guilty*, *sequential CM,* and *simultaneous CM*. The two CM groups were asked to execute a covert countermeasure before (*sequential*) or simultaneously with the “I saw it” button response, as detailed below.

The subjects were then presented with an inter-stimulus black screen for a variably random interval from 1400 ms to 1850 ms. A string of five identical numbers between 1 and 5 (e.g. 11111, 22222, etc.) was then presented to subjects for 300 ms. Subjects were asked to press the left button on the two button box using their index finger to all number strings (non-targets) except for the target string of ones. To targets, subjects were instructed to press the right button with their middle fingers. All number string display probabilities were comparable.

After consenting, subjects read written instructions for the practice task while the experimenter applied electrodes. The practice task was identical to the full task given to the *simple guilty* group, and included no *CM* instructions. Subjects were instructed to press the “I saw it” button as soon as they saw a date. Following the “I saw it” response, subjects did the target discrimination. The practice task continued until subjects completed 50 full trials.

Subjects then received instruction for their group and were asked to repeat the procedure in their own words. All subjects were instructed that the test was designed as a “mind reading or deception detection test” to discover concealed information. Subjects in the two CM groups were also told that they would be taught a method to “beat” the test. The instructions explained that subjects would be shown five dates. Once testing began, subjects completed 75 trials for each irrelevant date and the probe date, for a total of 375 trials (total number of programmed trials which yields approximately 50 trials per stimulus after artifacting) in 30 min. The ratio of probe to irrelevant trials was 1:4. At five random intervals throughout testing, the experimenter would stop the trial and ask the subject which date was last observed, in order to force attention to each presented date. Subjects who made more than one mistake were to be eliminated, but there were none. Errors of omission in “I saw it” responses were excluded from analysis. After completing the task, subjects were asked a series of debriefing questions including what they believed the purpose of the experiment was, and what strategies, if any, they used.

Subjects were divided into the following groups:

1. Simple guilty: A subject in this group was, on each trial, shown one of four irrelevant dates (irrelevant) or his/her birthday (probe). The instructions asked them to press the “I saw it” button as soon as possible after seeing the date. This group was designed to measure the effect of the probe without any countermeasure response to any irrelevant. Thus subjects were not guilty of a crime but of attempted concealment of birthdate recognition.
2. Sequential CM: Task instructions were identical to the *simple guilty* group EXCEPT that subjects in the *sequential CM* group were instructed to silently and mentally say either their first or their last names to two specifically assigned irrelevant dates and then respond with the “I saw it” response ‘as soon as possible’. Two, rather than all four, irrelevants were countered in order to avoid artificially enhancing probe P300 amplitudes due to probes being the only stimulus with no CM, (Meixner and Rosenfeld, 2010). Each name was assigned to a specific irrelevant.
3. Simultaneous CM: Task instructions were identical to *sequential CM* instructions EXCEPT that subjects in the *simultaneous CM* group were explicitly asked to press the “I saw it” button and make the covert CM response simultaneously.

Electrodes were placed at midline sites Fz, Cz, and Pz. Two linked reference electrodes were placed at the mastoids and a ground electrode was attached to the middle of the forehead. For all participants, electrode impedances were kept below 10 kΩ. EOG was recorded with two electrodes placed medially below and laterally above the right eye, this diagonal placement allowing for the tracking of eye movement and blinking. The criterion for artifact rejection varied given each subject's EOG artifact amplitudes but was always less than 50 μV. Trials in which EOG exceeded this amplitude were eliminated from analysis. Signals passed through amplifiers with 30 Hz low pass filter and .3 Hz high pass. Amplified output passed through a 16-bit Analog/Digital converter with sampling rate of 500 Hz.

We measured P300 amplitude using the peak–peak method as described in Soskins et al. (2001). This method is more sensitive to the detection of deception than the base-peak method used in earlier studies (Meijer, et al., 2007; Soskins, et al., 2001). The exact method has been detailed in many previous papers (Meijer, et al., 2007; Meixner and Rosenfeld, 2010; Rosenfeld and Labkovsky, 2010; Rosenfeld, et al., 2008; Rosenfeld, et al., 2004; Rosenfeld, et al., 2009; Soskins, et al., 2001; Winograd and Rosenfeld, 2011; Meixner and Rosenfeld, n.d). In this report, a window of 300–650 ms was used to find peak positivity and 650–1350 ms for peak negativity.

The bootstrap method (Wasserman and Bockenholt, 1986) was used to determine whether P300s for a given stimulus are greater than P300s for other stimuli. This process creates a distribution of probe-minus-irrelevant P300 differences for each subject. Past studies (Meijer, et al., 2007; Meixner and Rosenfeld, 2010; Rosenfeld and Labkovsky, 2010; Rosenfeld, et al., 2008; Rosenfeld, et al., 2004; Rosenfeld, et al., 2009; Soskins, et al., 2001; Winograd and Rosenfeld, 2011; Meixner and Rosenfeld, n.d) detailed the bootstrap method extensively. It registers a 'guilty' decision when a probe P300 is significantly larger than the P300 average of all irrelevants (Iall) on 90 out of 100 ($p<.1$) iterations of the bootstrap. We use here an identical criterion.

The mean "I saw it" reaction times to the probe, Iall, countered irrelevants (IC), and non-countered irrelevants (INC) for CM groups are given in Fig. 1a. Conditional probabilities of a target as opposed to a non-target stimulus following a probe or irrelevant was 1:4 regardless of which stimulus is presented. As shown in previous research (Rosenfeld et al., 2009) this 'symmetric' protocol does not affect behavioral or ERP results. Thus target versus non-target trial status was not used as a factor in analysis as probe RTs and ERPs followed by targets or non-targets were combined, and likewise with irrelevants. Using Probe and Iall RTs as the dependant variable, we conducted a 2 stimulus type (*probe, Iall*) × 3 group (*simple guilty, sequential CM, and simultaneous CM*) ANOVA. We found a significant main effect of group $F(2,68)=29.5$, $p<.001$, partial-$\eta^2=.465$. Post hoc Tukey tests found that the difference occurred between the *sequential CM* group and the other two groups ($p<.001$). We then performed a 1 × 3 ANOVA with Iall RT as the dependent variable and group (*simple guilty, sequential CM, and simultaneous CM*) as the factor, revealing significant differences among groups; $F(2,34)=19.6$, $p<.001$. Post hoc Tukey tests found that differences occurred between the *sequential CM* group (larger RTs) and the other two groups ($p<.001$). There was no difference in irrelevant RTs between *simple guilty* and *simultaneous CM* groups ($p=.748$).

To explore RT differences between responding simultaneously and sequentially in more detail a 2 group (*simultaneous CM versus sequential CM*) × 3 stimulus type (*probe, IC, and INC*) ANOVA was conducted, identifying a significant main effect of Group; $F(1,69)=55.9$, $p<.001$, $\eta_p^2=.448$. The *sequential CM* reaction times were significantly longer than the *simultaneous CM* RTs. As expected, there was also a significant main effect of Stimulus Type; $F(2,69)=10.2$, $p<.001$, $\eta_p^2=.229$. Additionally there was a significant interaction of Group and Stimulus type; $F(2,69)=5.65$, $p<.01$, $\eta_p^2=.141$.

In further investigating this interaction, separate 1 × 3 ANOVAs were conducted on both the *simultaneous CM* and *sequential CM* groups with RT as the dependant variable and stimulus type (*probe, IC, and INC*) as the factor. The ANOVA on the *simultaneous CM* RT data showed that there were no significant differences between reaction times for probe, countered, and non-countered irrelevant stimuli; $F(2,33)=.477$, $p=.625$. In contrast, the ANOVA conducted on the *sequential CM* data showed that there were significant differences in RT for the different stimuli types; $F(2,33)=14.0$, $p=.001$. Post hoc Tukey tests on the *sequential CM* ANOVA identified differences between countered and non-countered RTs ($p<.001$), and between countered and probe RTs ($p<.001$), while no difference between non-countered and probe RTs was detected ($p=.749$).

To examine differences in P300 ERPs, a 2 stimulus type (*probe, Iall*) × 3 group (*simple guilty, sequential CM, and simultaneous CM*) ANOVA was conducted with P300 amplitude as the dependant variable. These data are shown in Fig. 1b. As expected, a main effect of Stimulus Type was observed; $F(1,68)=18.3$, $p<.001$, $\eta_p^2=.212$. Neither the effect of group nor the interaction of Group x Stimulus Type were significant. To analyze these differences in further detail, a 3 stimulus type (*probe, IC, and INC*) × 2 group (*sequential CM and simultaneous CM*) ANOVA was conducted on the P300 ERPs of the *simultaneous CM* and *sequential CM* groups. As expected, a main effect of Stimulus Type was observed; $F(2,69)=6.86$, $p=.002$, $\eta_p^2=.166$. Both the effects of Group, $F(1,69)=.456$, $p=.502$, and the interaction of Group and Stimulus Type, $F(2,69)=.037$, $p=.963$, were insignificant. Post hoc Tukey tests showed that only the probe and non-countered P300s differed significantly ($p<.001$) with probe P300s being larger. Both the amplitude differences between countered and non-countered irrelevant P300s ($p=.121$) and between countered irrelevant and probe P300s ($p=.208$) were insignificant, although the results suggest a trend that countered irrelevant P300s were larger than non-countered irrelevant P300s.

Detection rates for probe versus Iall P300s within subjects appear to be similar across groups at 10/12 for *simple guilty*, 11/12 for *sequential CM*, and 11/13 for *simultaneous CM*.

As feared, we observed that with 4 total irrelevants, performing a CM and making the "I saw it" response simultaneously eliminates the ability to use within-block RT of that response as an indicator of CM use, since this *simultaneous CM* style equalizes RTs to both probe and

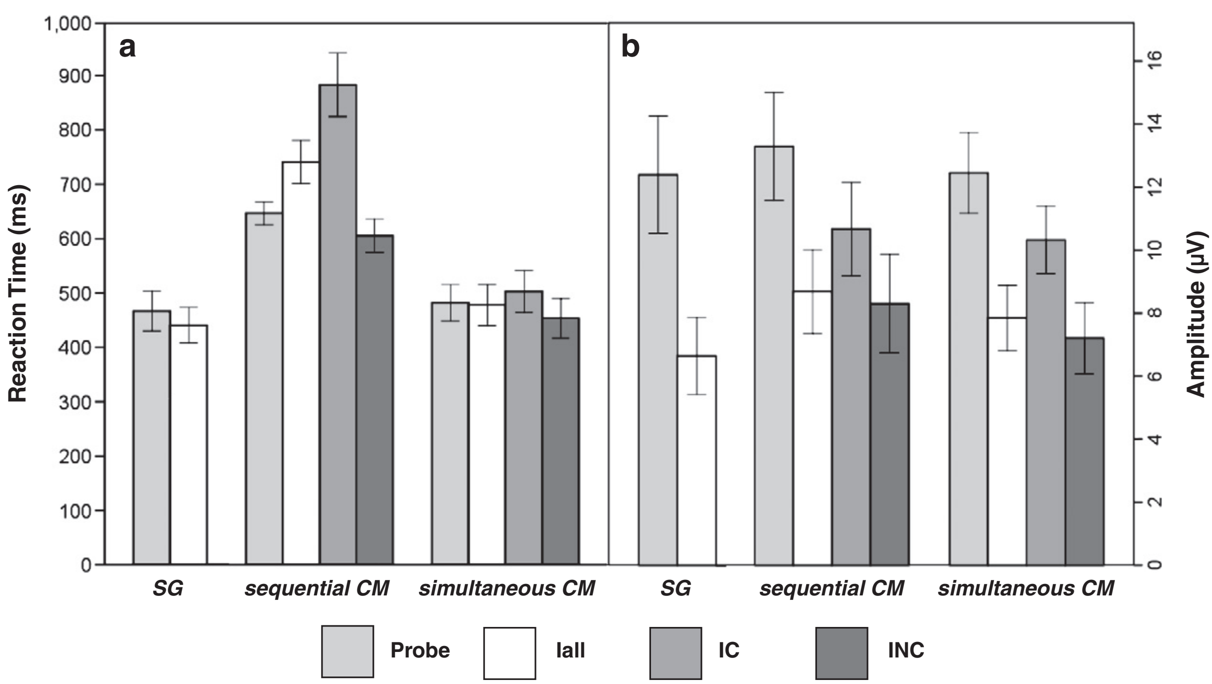

Fig. 1. Probe, Iall, IC, and INC reaction times (a) and peak–peak P300 ***sequential CM*** (b) in the simple guilty (SG), sequential CM, and simultaneous CM groups. Due to P300 latency jitter differences, Iall amplitude calculated on about 200 ERPs does not fall midway between IC and INC amplitudes, each calculated here by averaging 2 irrelevant averages, each averages of 50 ERPs.

irrelevant stimuli. On the other hand, this CM style does not affect the diagnostic ability of the CTP based on P300 amplitude differences between probe and irrelevant P300 averages within individuals. However, since the presence of elevated RT levels to specific irrelevants can be indicative of CM use, a new strategy for eliciting an RT elevation should be searched for, possibly through increasing the number of irrelevant items and the consequent difficulty of the task.

References

Ben-Shakhar, G., Dolev, K., 1996. Psychophysiological detection through the guilty knowledge technique: the effects of mental countermeasures. Journal of Applied Psychology 81, 273–281.

Ben-Shakhar, G., Elad, E., 2002. Effects of questions' repetition and variation on the efficiency of the guilty knowledge test: a reexamination. Journal of Applied Psychology 87, 972–977.

Committee to Review the Scientific Evidence on the Polygraph, National Research Council, 2003. The Polygraph and Lie Detection. National Academies Press, Washington DC.

Elaad, E., Ben-Shakhar, G., 1997. Effects of item repetitions and variations on the efficiency of the Guilty Knowledge Test. Psychophysiology 34, 587–596.

Farwell, L., Donchin, E., 1991. The truth will out: interrogative polygraphy ("Lie detection") with event-related brain potentials. Psychophysiology 28 (5), 531–547.

Honts, C., Devitt, M., Winbush, M., Kircher, J., 1996. Mental and physical contermeasures reduce the accuracy of the concealed knowledge test. Psychophysiology 33, 84–92.

Honts, C., Amato, S., Gordon, A., 2001. Effects of spontaneous countermeasures used against the comparison question test. Polygraph 30 (1), 1–9.

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

Meijer, E., Smulders, F.T., Merckelbach, H., Wolf, A., 2007. The P300 is sensitive to concealed face recognition. Int J Psychophysiology 66 (3), 231–237.

Meixner, J., Rosenfeld, J., 2010. Countermeasure mechanisms in a P300-based Concealed Information Test. Psychophysiology 47 (1), 57–65.

Meixner, J., & Rosenfeld, J. (n.d.). A Mock Terrorism Application of the P300-Based Concealed Information Test. *Psychophysiology* .

Mertens, R., Allen, J., 2008. The role of psychophysiology in forensic assessments: deception detection, ERPs, and virtual reality mock crime scenarios. Psychophysiology 45 (2), 286–298.

Rosenfeld, J., Labkovsky, E., 2010 published online. 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.

Rosenfeld, J., Angell, A., Johnson, M., Quian, J., 1991. An ERP-based, control-question lie detector analog: algorithms for discriminating effects within individual's average waveforms. Psychophysiology 18, 319–335.

Rosenfeld, J., Soskins, M., Bosh, G., Ryan, A., 2004. Simple effective countermeasures to P300-bases tests of detection of concealed information. Psychophysiology 41 (2), 205–219.

Rosenfeld, J., Labkovsky, E., Winograd, M., Lui, M., 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.

Rosenfeld, J., Tang, M., Meixner, J., Winograd, M., Labkovsky, E., 2009. The effects of asymmetric vs. symmetric probability of targets following probe and irrelevant stimuli in the complex trial protocol for detection of concealed information with P300. Physiology & Behavior 98, 10–16.

Soskins, M., Rosenfeld, J.P., Niendam, T., 2001. The case for peak-to-peak measurement of P300 recorded at 3 Hz high pass filter settings in detection of deception. Int. J. Psychophysiology 40 (2), 173–180.

Wasserman, S., Bockenholt, U., 1986. Bootstrapping: applications to psychophysiology. Psychophysiology 26 (2), 208–221.

Winograd, M.R., Rosenfeld, J.P., 2011. Mock crime application of the Complex Trial Protocol (CTP) P300-based concealed information test. Psychophysiology 48, 155–161.