Brain and Cognition 90 (2014) 142–150

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Detecting concealed information using feedback related event-related brain potentials

Liyang Sai[a,b,1], Xiaohong Lin[a,d,1], Xiaoqing Hu[c,*], Genyue Fu[a,*]

[a] Department of Psychology, Zhejiang Normal University, Jinhua, China
[b] School of Psychology & Cognitive Science, East China Normal University, Shanghai, China
[c] Department of Psychology, Northwestern University, Evanston, USA
[d] Bioimaging Core, Faculty of Health Sciences, University of Macau, Taipa, Macau, Macau SAR, China

ARTICLE INFO

Article history:
Accepted 20 June 2014

Keywords:
Conceal information test
Feedback processing
Deception detection
Event-related potentials

ABSTRACT

Employing an event-related potential (ERP)-based concealed information test (CIT), the present study investigated (1) the neurocognitive processes when people received feedbacks regarding their deceptive/truthful responses and (2) whether such feedback-related ERP activities can be used to detect concealed information above and beyond the recognition-related P300. During the CIT, participants were presented with rare, meaningful probes (their own names) embedded within a series of frequent yet meaningless irrelevants (others' names). Participants were instructed to deny their recognition of the probes. Critically, following participants' responses, they were provided with feedbacks regarding whether they succeeded or failed in the CIT. Replicating previous ERP-based CITs, we found a larger P300 elicited by probe compared to irrelevant. Regarding feedback-related ERPs, a temporospatial Principle Component Analyses found two ERP components that were not only sensitive to feedback manipulations but also can discriminate probe from irrelevant: an earlier, central-distributed positivity that was elicited by "success" feedbacks peaked around 219 ms; and a later, right central-distributed positivity that was also elicited by "success" feedbacks, peaked around 400 ms. Importantly, the feedback ERPs were not correlated with P300 that was elicited by probe/irrelevant, suggesting that these two ERPs reflect independent processes underlying memory concealment. These findings illustrate the feasibility and promise of using feedback-related ERPs to detect concealed memory and thus deception.

1. Introduction

Researchers have devoted great efforts to the development of event-related potential (ERP)-based concealed information tests (CITs). A majority of these studies focused on the P300, a large, positive deflection of brainwaves that occurs between 300 and 800 ms after stimulus onset. P300 is sensitive to a range of factors such as subjective probability, task-relevance and available cognitive resources (Donchin & Coles, 1988; Johnson, 1988). These characteristics of P300 have been employed in P300-based CITs that aim to identify whether the examinee recognizes the concealed information or not, regardless of his or her verbal report (Rosenfeld, 2011; Rosenfeld, Hu, Labkovsky, Meixner, & Winograd, 2013). Specifically, examiners compare the P300s elicited by two types of information in the CIT: a rarely presented, crime-related information item (e.g. the weapon used in a murder, also referred to as a "probe" item) and a series of crime-irrelevant alternatives (e.g. other weapons that were not used in the murder, also referred to as irrelevant items). If the probe is associated with significantly larger P300 than the irrelevant, then a recognition diagnosis is made. If, however, no systematic difference is found between the probe and irrelevants, then a non-recognition diagnosis is made. Indeed, it has been found that considerable P300s can be elicited by a range of stimuli, including incidentally acquired crime-related information as well as well-rehearsed personal information (e.g. one's hometown or first name probe items) (Hu, Hegeman, Landry, & Rosenfeld, 2012; Hu, Pornpattananangkul, & Rosenfeld, 2013; Rosenfeld et al., 2008). The fundamental mechanism underlying P300-based CITs is detection of the memory status of the information of interest (i.e. old vs. new; recognized vs. not recognized).

* Corresponding authors. Address: Department of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208, USA (X. Hu). Address: Department of Psychology, Zhejiang Normal University, Yingbin Rd., Jinhua 321004, China (G. Fu).

E-mail addresses: xhu@u.northwestern.edu (X. Hu), fugy@zjnu.cn (G. Fu).

[1] These authors contributed equally to this work

http://dx.doi.org/10.1016/j.bandc.2014.06.012
0278-2626/

In addition to the memory mechanism underlying the P300-based CIT, the influence of intentional concealment on ERPs (e.g. P300) in a CIT has also been investigated. The central aim of these studies is two-fold: (1) from a theoretical view, to explore whether the concealment intention or deceptive response involve mechanisms that are independent of the memory/recognition account underlying CIT; and (2) from an applied view, to determine whether an intention to conceal information and deceptive responses can influence the detection efficiency of ERP-based CIT (Kubo & Nittono, 2009; Rosenfeld, Hu, & Pederson, 2012; Verschuere, Rosenfeld, Winograd, Labkovsky, & Wiersema, 2009). For instance, in Verschuere et al. (2009) and Rosenfeld et al. (2012), researchers examined two groups: a deception group and a control group. Participants in the deceptive group were instructed to conceal probe items while participants in the control group were told to perform a target/no target discrimination task: thus in this group, no deception or concealment was mentioned (Rosenfeld et al.,2012; Verschuere et al., 2009). However, results are inconsistent across studies: whereas Verschuere et al. (2009) and Rosenfeld et al. (2012) found that deception did improve individual detection efficiency of the CIT, results from other studies suggest that an intention to conceal does not modulate the P300s in the CIT (Kubo & Nittono, 2009). Recently, Rosenfeld et al. (2012) added a novel manipulation to investigate the role of deception in the P300-based CIT. Specifically, in addition to the instruction that explicitly required participants to respond deceptively to probe items, the study also included periodic feedbacks to maintain participants' awareness that they were giving deceptive responses to probes. Results showed that when deceptive participants received periodic feedback regarding their deception, they showed larger P300 amplitudes than participants in the control group who had no intention to conceal the information (see also, Hu et al., 2013). Because a previous study that manipulated only instruction failed to find enhanced P300s among the deceptive participants, the enlarged P300 responses observed in Rosenfeld et al. (2012) was ascribed specifically to the use of periodic feedback that reminded participants of their deceptive responses. This feedback manipulation was recently applied in the complex trial protocol (Rosenfeld et al., 2008) and replicated the effect that such feedback can enhance the detection efficiency based on P300 (Hu et al., 2013). Moreover, receiving feedback regarding information concealment elicited higher frontal–central negativities between 200 and 400 ms, suggesting the involvement of performance monitoring processes during information concealment or deception (Gamer & Berti, 2010; Hu et al., 2013). Thus, from an applied view, it seems that an intention to conceal, especially when feedbacks are used to emphasize one's deceptive responses throughout the test, can improve the detection efficiency of ERP-based CIT (Hu et al., 2013; Matsuda et al., 2013; Rosenfeld et al., 2012). However, the neurocognitive processes associated with feedback processing among those who intentionally conceal information are still unclear, and it remains to be explored that whether ERP activities during feedback can discriminate probe information from irrelevant information. These will be the major questions we aim to explore in the present study.

Here, we aim to examine ERPs that are directly elicited by feedbacks during a CIT to investigate the neurocognitive processes underlying feedback processing. Previous studies in feedback processing consistently find a negative deflection of ERPs between the 200 and 300 ms time window that is sensitive to negative feedbacks in comparison with positive feedbacks, which is termed the Feedback-negativity (FN, also known as feedback-related negativity FRN, or feedback error-related negativity fERN (e.g. Holroyd, Larsen, & Cohen, 2004). Such negative feedbacks are usually contingent upon participants' performance or choices, such as incorrect motor responses (Miltner, Braun, & Coles, 1997), monetary loss (Gehring & Willoughby, 2002) and unexpected outcomes (Ferdinand, Mecklinger, Kray, & Gehring, 2012; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003). The amplitude of the FN is suggested to reflect the difference between actual and expected outcomes (Holroyd & Coles, 2002; Nieuwenhuis, Holroyd, Mol, & Coles, 2004), or participants' evaluation of the motivational impact of ongoing events (Gehring & Willoughby, 2002; Yeung, Holroyd, & Cohen, 2005).

Here we adapted to a CIT that gave feedbacks following each test item (e.g. probe and irrelevant). As in previous CIT studies, participants were presented with a rare probe item and a series of irrelevant items. Moreover, they were instructed to deny the knowledge of probe items via button pressing (e.g. Rosenfeld et al., 2012; Verschuere et al., 2009). Critically, after participants made a button press to the CIT stimulus, we provided them with feedbacks regarding whether they had successfully deceived the brainwave-based lie detector. Here, the feedbacks were given randomly and were not contingent upon their behavior. Because of the great motivational significance of probe to participants, e.g. they need to try to conceal the probe and avoid being detected; we predicted that feedbacks following probes would elicit larger feedback-related potentials than irrelevant (Luo, Sun, Mai, Gu, & Zhang, 2011; Yeung et al., 2005).

To fully explore the feedback-related ERPs, we used a temporo-spatial Principle Component Analysis (PCA) to quantify the feedback-related ERPs. As a data-driven approach, PCA has been widely used in ERP research to decompose raw ERPs components along temporal and spatial domain (Donchin & Heffley, 1979; Spencer, Dien, & Donchin, 2001). An advantage of PCA is that it allows researchers to separate ERPs activities that may overlap with each other in time/space. In particular, PCA has been used to quantify feedback-related ERPs in monetary feedback processing (Foti, Weinberg, Dien, & Hajcak, 2011), as well as memory-related ERPs in the P300-based CITs (Lui & Rosenfeld, 2009). Because the present study aims to explore the feedback processing during information concealment, the PCA will be particularly useful to isolate ERP-of-interest that is sensitive to our independent variables: feedback valence and stimulus type.

Finally, we predict that as the feedback-related ERPs reflect participants' motivational process to evaluate whether their behavior/responses is success or not, this ERP pattern should be independent of the P300 that mainly reflects memory processes such as item recognition. Such independence information would be valuable from an applied perspective, as this suggests that the feedback-related ERPs can identify concealed information above and beyond the P300.

2. Methods

2.1. Participants

Twenty participants were recruited (all males, M_{age} = 21.6 years, SD = 2.7 years), three of which were excluded from ERP analyses due to excessive artifacts. All participants had normal or corrected to normal vision, and were right-handed. None had a history of any neurological or psychiatric disorders. The study was approved by ethics committee of Zhejiang Normal University.

2.2. Procedures

Upon entering the laboratory, each participant signed an informed consent form. For all participants, the target was a Chinese celebrity's name "Liu Dehua," and the probe was their own name. Four irrelevants were selected from a list of ordinary Chinese names. Before experiment, a questionnaire was conducted to make sure the irrelevant names do not have special meaning for the participants.

As electrodes were applied, an experimenter read the instructions as follows: "Now, imagine that you are a spy who is arrested by the police. You have to prove your innocence by beating the lie detector. In the lie detecting test, you are going to see a series of names on the display. And you are to press the "F" which means, "I recognize the name" when you see the target name "Liu Dehua". You will therefore be telling the truth since you do recognize it as your target name. Otherwise, you press the "J" which means "I don't recognize the name" to all other names that are not targets. But one of these non-target names you will see is your own name. When you press the "I don't recognize it" button for your name you will be lying. You don't recognize it as your target, but you do recognize it as your own name. Then, the lie detector will detect whether you are lying or not. You can see "detecting" on the display screen. And then the lie detector will give you a feedback for either "truth" or "lie" based on the analyses of your brain waves". In fact, the feedbacks were presented randomly.

Participants were seated about 1 m in front of the computer. Each trial began with a 500-ms fixation point. Then a name was presented for 300 ms. When the stimulus appeared, the participants were instructed to press one of two buttons as quickly and accurately as possible. After a 1000 ms blank, subject would see "detecting" on the screen for 500 ± 100 ms, which meant the lie detector was detecting. Finally, the feedback to the detecting was presented as "truth" or "lie" for 1000 ms following a 500 ms blank (see Fig. 1). Thus, there were four conditions in the feedback stage: probe-truth (success); probe-lie (failure); irrelevant-truth (success); irrelevant-lie (failure). The probe and each irrelevant repeated 80 times, and half of them were followed by the feedback of "truth", and the other half of the stimuli were followed by the feedback of "lie". The target was also repeated 80 times, but target feedback was based on participant performance (i.e. the feedback would be "truth" if the participant responded correctly, and the feedback would be "lie" if the participant responded incorrectly. Thus, there were $6 \times 80 = 480$ trials total. Every 30 trials (about 2 min), participants were allowed to take a break. The whole experiment lasted about 30 min.

At the end of the experiment, we asked participants to rate their subjective feelings for the four types of feedbacks: probe-truth (success); probe-lie (failure); irrelevant-truth (success); irrelevant-lie (failure) on a five-point scale (1 = very upset, 5 = very excited).

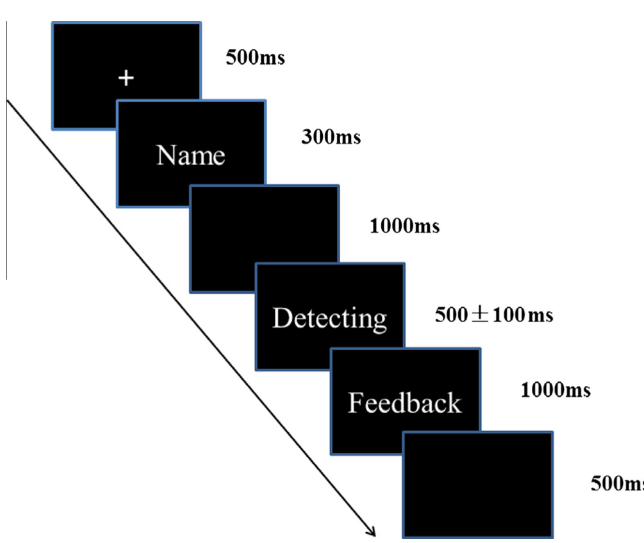

Fig. 1. Task structure.

2.3. EEG acquisition

Continuous EEGs were recorded from 32 scalp sites using Ag/AgCl electrodes embedded in an elastic cap (Neuroscan Inc., USA) according to the international 10–20 system. On-line recordings were referenced on the left and right mastoids. Electrode impedances were kept below 5 kΩ. The vertical electro-oculograms (EOGs) were recorded above and below the right eye; the horizontal EOGs were recorded from electrodes placed at the outer canthus of left eye and right eye. The sampling rate was set to 1000 Hz.

For offline analyses, continuous EEGs were first filtered with a 30 Hz low-pass filter. Continuous EEGs were then segmented and locked to the CIT stimuli and the feedback stimuli, respectively. For the CIT stimuli (e.g. probe or irrelevant), each epoch contained a 200 ms pre-stimulus baseline and a 1000 ms time window after stimuli presentation. For feedback stimuli (e.g. success vs. fail), each epoch contained a 200 ms pre-stimulus baseline and a 1000 ms time window after feedback stimuli presentation. Trials exceeding ±80 μV were defined as artifacts and were excluded from averaging. For ERPs locked to CIT stimuli, we focused on the P300, the amplitude of which was calculated as the mean of the maximal 100-ms segment between 300 and 800 ms after probe/irrelevant at Pz, which is used in many previous P300-based CIT study (e.g. Hu & Rosenfeld, 2012; Rosenfeld et al., 2013; Soskins, Rosenfeld, & Niendam, 2001). ERPs locked to feedbacks were quantified using temporospatial PCA analyses. Temporospatial PCA extracts linear combinations of data points that meet certain criteria that tend to distinguish between consistent patterns of electrocortical activity (cf. Foti et al., 2011). This analysis was conducted using the ERP PCA Toolbox (Version 210; (Dien, 2012)). A temporal PCA was first performed on the data, using 1200 time points (1000 samples multiplied by one trial-plus-baseline length of 1200 ms) per trial as variables and participants, recording sites, and trial types as observations. The temporal PCA yielded 13 factors based on the resulting scree plot. They were submitted to Promax rotation (Dien, Khoe, & Mangun, 2007). Then a spatial PCA was conducted with electrodes[1] as variables and all participants, conditions, and temporal factor scores as observations. Infomax rotation was used to rotate to independence in the spatial dimension. Two spatial factors were extracted for each temporal factor, yielding a total of 26 temporospatial factor combinations. Of these, 11 factors that each accounted for more than 1% of the variance were retained for further examination. Factors of interest were scored using the peak values of the virtual component. Analyses of variance (ANOVAs) were conducted by using SPSS 20.0. The Greenhouse-Geisser correction was applied when the assumption of sphericity was violated. Post hoc comparisons were computed with Fisher's Protected Least Significant Difference.

3. Results

3.1. Behavioral results

Sixteen participants' behavioral data were included for analysis because one participant's behavioral data was missing. A paired sample t-test showed that reaction time (RT) on probe trials ($Mean$ = 599.9 ms, Standard Deviation SD = 30.675) were significantly longer than that on irrelevant trials (M = 544.5 ms, SD = 31.725) ($t(15) = 3.6$, $p < 0.01$). For accuracy of probe and irrelevants, a paired sample t-test showed that there was no significant difference between the accuracy on probe trials (M = 0.98,

$SD = 0.053$) and irrelevant trials ($M = 0.99$, $SD = 0.0325$) ($t(15) = 1.4$, $p > 0.05$).

3.2. ERP results

3.2.1. P300 to probe vs. irrelevant

Based on previous studies, we focused our P300 analyses at Pz (e.g. Rosenfeld et al., 2013). A paired sample *t*-test was conducted on amplitude of P300 (a maximal positive 100-ms segment average between 300 and 800 ms after probe or irrelevant). The results revealed that the probe ($M = 12.56$ μv, $SD = 3.71$) elicited a larger P300 than irrelevants ($M = 6.77$ μV, $SD = 3.20$) ($t(16) = 7.73$, $p < 0.01$). (For the ERP waveform and Topography, see Fig. 2 and 3).

3.2.2. The ERPs during the feedback stage

In correspondence with previous studies focusing on FN and the following P300 (Holroyd et al., 2004; Yeung & Sanfey, 2004), we chose four factors for analyses based on their temporal and spatial similarities with FN and the feedback-related P300. These factors were: a positivity peaking at 219 ms at Cz; a positivity peaking at 303 ms at Cz; a positivity peaking at 400 ms at C4; and a positivity peaking at 453 ms at CPz. We conducted 2 (stimulus type: probe, irrelevant) by 2 (feedback: success, failure) repeated measure ANOVAs based on factor scores from the four temporospatial factors of interest (for a summary, see Table 1; for ERP waveforms and Topographies, see Figs. 2–6).

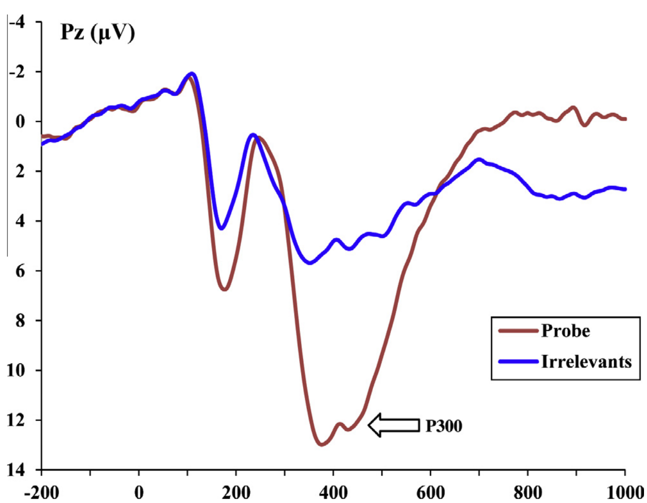

Fig. 2. Grand average ERPs at Pz during the stage of making response to probe or irrelevants.

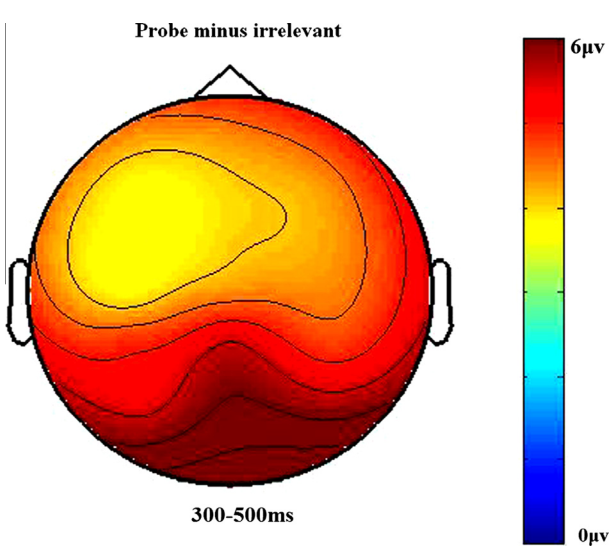

Fig. 3. Scalp topography of the difference wave between ERP responses that are locked to probe and irrelevants.

3.2.3. The central positivity at 219 ms

A 2 (stimulus probe vs. irrelevant) by 2 (feedback: success vs. failure) within-subject repeated measure ANOVA with the 219 positivity at Cz revealed a significant main effect of stimulus type, $F(1, 16) = 28.375$, $p < 0.001$, $\eta^2 = .639$, with a more positive amplitude following feedbacks regarding probe items ($M = 4.74$ μV, $SD = 0.55$) compared to feedbacks regarding irrelevant items ($M = 3.36$ μV, $SD = 0.41$). A significant main effect of feedback was also found, $F(1,16) = 22.603$, $p < 0.001$, $\eta^2 = .586$, with more positive following success ($M = 4.40$ μV, $SD = 0.49$) than following failure ($M = 3.70$ μV, $SD = 0.47$). No significant interaction was found.

3.2.4. The central positivity at 303 ms

The same ANOVA on the 303 ms positivity at Cz revealed that there was a significant main effect of stimulus type, $F(1,16) = 16.5$, $p < 0.001$, $\eta^2 = .508$, with a higher positivity following feedbacks regarding probe items ($M = 3.01$ μV, $SD = 0.64$) than feedbacks regarding irrelevant items ($M = 1.05$ μV, $SD = 0.35$). There was no other significant main effect or interaction.

3.2.5. The right-central positivity at 400 ms

The same ANOVA with the 400 ms positivity at C4 revealed a significant main effect of stimulus type, $F(1,16) = 12.367$, $p = 0.003$, $\eta^2 = .436$, as feedbacks following the probe elicited a more positive amplitude ($M = 2.33$ μV, $SD = 0.51$) than that elicited by feedbacks following irrelevant stimuli ($M = 1.13$ μV, $SD = 0.33$). In addition, a significant main effect of feedback was also found, $F(1,16) = 36.825$, $p < 0.001$, $\eta^2 = .697$, with success feedbacks eliciting more positive amplitude ($M = 2.21$ μV, $SD = 0.43$) than failure feedbacks ($M = 1.25$ μV, $SD = 0.37$). No interaction was found between these two factors.

3.2.6. The centroparietal positivity at 453 ms

The same ANOVA with the 453 ms positivity at CPz revealed a significant main effect of stimulus type, $F(1,16) = 42.11$, $p < 0.001$, $\eta^2 = .725$, as the amplitude associated with feedbacks following the probe ($M = 5.09$ μV, $SD = 0.76$) was more positive than the amplitude associated with feedbacks following irrelevant stimuli ($M = 1.95$ μV, $SD = 0.56$). There was no other significant main effect or interaction.

3.3. Participants' subjective feelings following feedbacks

A 2 (stimulus type: probe, irrelevant) by 2 (feedback valence: success, failure) within-subject repeated measurement ANOVA was conducted with the participants' subjective ratings. The results showed that there was a significant main effect of stimulus type, $F(1,16) = 10.2$, $p < .01$, $\eta^2 = .34$, with a higher level of upset the participants reported when receiving feedbacks following probes ($M = 2.97$) than feedbacks following irrelevants ($M = 3.36$). A significant main effect of feedback valence was also found, $F(1,16) = 45.4$, $p < .001$, $\eta^2 = .69$, with higher level of upset following failure ($M = 2.5$) than success feedbacks ($M = 3.8$). There was no significant interaction between stimulus type and feedback valence.

Table 1
Description and analysis of variance results for each temporospatial factor.

Temporospatial factor	Peak loading (ms)	Variance (%)	Polarity	Spatial distribution	Main effect of stimulus type $F(\eta^2)$	Main effect of feedback $F(\eta^2)$	Stimulus type* feedback F
TF1SF1	453	3.9	+	CPz	42.11*** (0.73)	2.44	2.00
TF3SF1	219	8.9	+	Cz	28.36***(0.64)	22.60***(0.59)	2.44
TF4SF1	303	5.5	+	Cz	16.5***(0.51)	<1	<1
TF6SF1	400	3.1	+	C4	12.36**(0.46)	36.83***(0.70)	<1

Note: df = 1,16.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

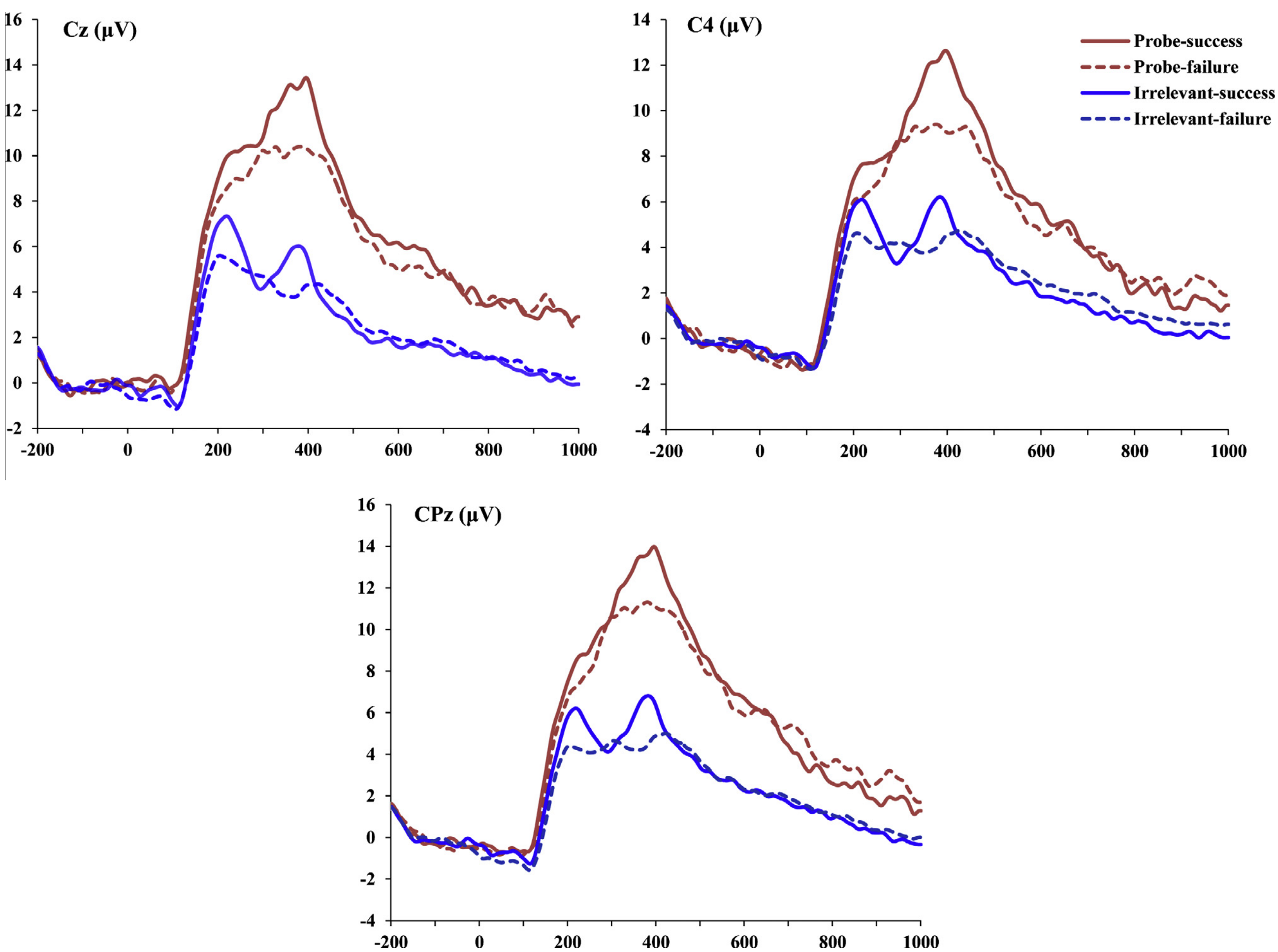

Fig. 4. Original ERPs locked to feedbacks (success vs. fail) following probe or irrelevant.

3.4. Independence testing for recognition-related P300 and feedback-related ERPs

We predict that as the probe-related P300 mainly indicates recognition and memory processes in the CIT, and the feedback-related ERP mainly reflects motivational processes such as outcome evaluation, these two ERPs may not correlate with each other. To formally test this hypothesis, we ran a correlation analyses between probe/irrelevant P300s and the four feedback-related ERP components that were derived from the PCA. Results showed that across participants, there was no correlation between P300 and any of the four feedback-related ERP components, indeed the correlation was almost negligible (see Table 2). This pattern of results supports our prediction that these two ERPs may reflect different, independent psychological processes underlying memory concealment.

3.5. Receiver Operating Characteristic (ROC) analysis

Receiver Operating Characteristic (ROC) Analysis were conducted based on a comparison between the guilty group ran here with a simulated innocent group ($N = 17$) (For details, see Carmel, Dayan, Naveh, Raveh, & Ben-Shakhar, 2003; Meijer, Smulders, Johnston, & Merckelbach, 2007). Such a group was created by randomly drawing value from a standard normal distribution (mean = 0, standard deviation = 1). Specifically, for feedback stage, ten values were randomly drawn from a standard normal distribution. Each value represents one of ten conditions in feedback stage

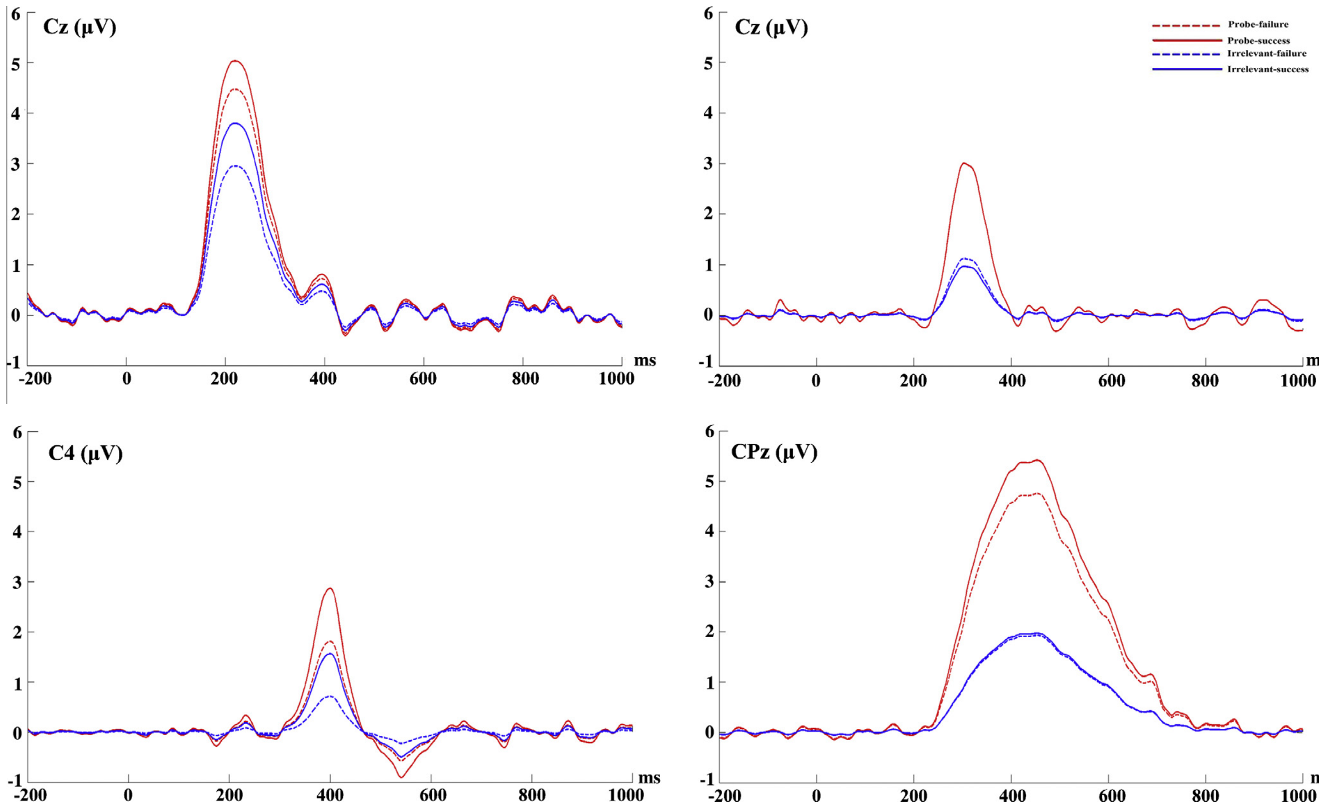

Fig. 5. Virtual ERP components extracted from the temporospatial PCA.

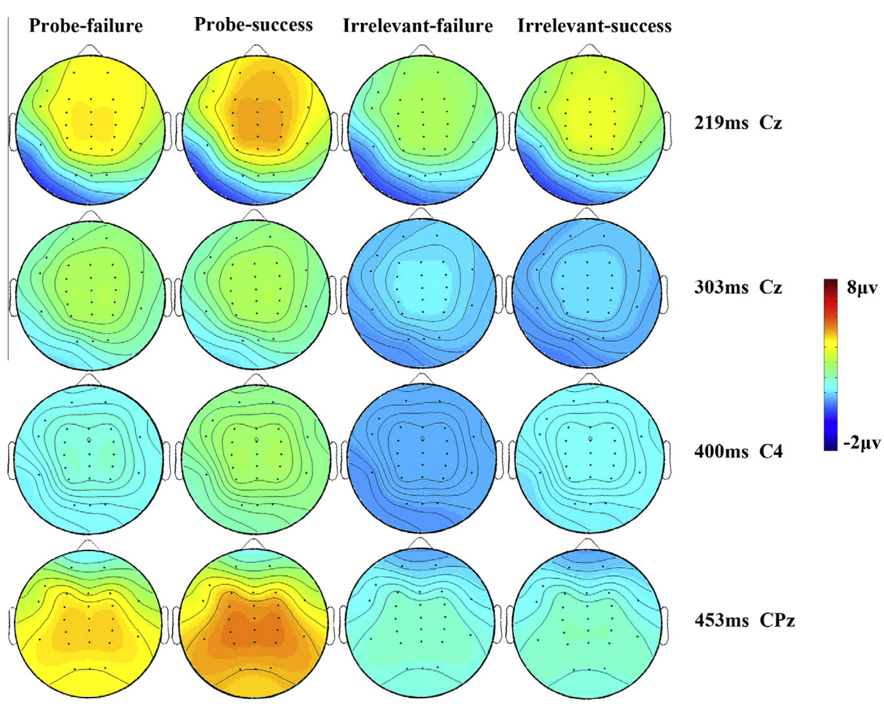

Fig. 6. Scalp topographies of temporospatial Principle Component Analyzed-ERPs locked to feedbacks (success vs. fail) following probe or irrelevant.

(probe-success vs. failure, irrelevant1, 2, 3, 4-success vs. failure,). This process was repeated 40 times (because each condition in the guilty group was repeated 40 times). These forty values for each condition were averaged to represent a score of each condition for one innocent participant. For recognition-related P300 in the CIT, five values were randomly drawn from a standard normal distribution. Each value represents one of five conditions (probe, irrelevant1-4). This process was repeated 80 times (because both probe and irrelevants were repeated 80 times in the guilty group). These procedures were repeated 17 times to obtain a distribution

Table 2
Correlations between RT, P300 and feedback-related activities.

	RT	P300	The central positivity at 219 ms	The central positivity at 303 ms	The right-central positivity at 400 ms
RT					
P300	−.007				
The central positivity at 219 ms	−.28	.003			
The central positivity at 303 ms	−.32	−.16	.72**		
The right-central positivity at 400 ms	−.09	−.24	.65**	.87**	
The centroparietal positivity at 453 ms	−.17	−.13	.72**	.75**	.79**

** $p < .01$.

that consisted of 17 innocent participants. Based on these procedures, we created the values of probe and irrelevants for P300 during CIT, and the feedback-related ERP components at 219 ms and 400 ms during feedback stage for this simulated innocent population.

ROC analyses were conducted based on probe-minus-irrelevants (four irrelevants were averaged) of P300, the components in 219 ms and 400 ms during feedback. The results showed that all the three components can discriminate guilty from innocent group above chance (95% CIs were given in parentheses) P300: AUC = 1.00 (1.00, 1.00), $p < 0.01$; P219: AUC = 0.94 (0.83, 1.00), $p < 0.01$); P400: AUC = 0.84 (0.69–0.99), $p < 0.01$).

4. Discussion

In the present study, we examined (1) the neurocognitive processes when participants received feedbacks about their deceptive/truthful responses; and (2) whether such feedback-related ERPs can discriminate between probe and irrelevant items in an ERP-based CIT. As the P300s to the probe are significantly larger than the irrelevant stimuli, replicating the most basic findings in P300-based CIT, we focus our discussion on the novel finding: the feedback-related ERPs. Specifically, using a temporospatial PCA method, we found that a series of ERPs that were sensitive to feedback valence as well as stimulus type. Specifically, we found: a central-distributed positivity that peaked at 219 ms and a right central-distributed positivity that peaked at 400 ms, both of which were sensitive to feedback valence (success vs. fail) and stimulus type (probe vs. irrelevant). Moreover, we found a central-distributed positivity that peaked at 303 ms and a centroparietal-distributed positivity that peaked at 453 ms, both of which were sensitive to stimulus type. Critically, neither of these feedback-related ERPs was correlated with P300 that was elicited by probe or irrelevant, suggesting feedback processes and item recognition played different role in memory concealment.

First, we found that the central positivity that peaked at 219 ms was larger following success than failure, suggesting that this positivity was sensitive to feedback valence. It should be noted that our PCA analyses found a larger positivity for success than for failures, which seems at odds with previous studies that usually found a negative deflection following negative feedback than positive feedback. This feedback-related negativity is typically understood to be a negative deflection that is elicited by negative feedback stimuli, and this negative deflection is absent following positive feedback (e.g. Gehring & Willoughby, 2002). However, recent studies that employed PCA found that positive feedbacks elicited a positive-going deflection in the time range of feedback-related negativity. These findings suggest that the FN is actually driven by a positive deflection at the frontocentral regions that is sensitive to gains compared to losses (Foti et al., 2011; Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Holroyd et al., 2003; Lole, Gonsalvez, Barry, & De Blasio, 2013). Unlike the conventional difference wave approach, these studies applied temporospatial PCA to measure the FN with the goal of distinguishing FN from later, usually larger overlapping ERP activities such as the P300 (e.g. Foti et al., 2011; Lole et al., 2013). In the current study, we also used a temporospatial PCA to examine the latent spatial and temporal characteristics of feedback-related brain potentials. We found the positivity that peaked at 219 ms was larger following success than failure. Given this positivity's sensitivity to feedback valence, and given the scalp distribution and its temporal feature, this activity may mirror the FN.

Importantly, our results also showed that this feedback-related positivity (FRP) can discriminate concealed personal information from irrelevant items, as feedbacks regarding participants' own name (probe) elicited a more positive FRP than feedbacks regarding others' names (irrelevant items). This result demonstrated that the FRP have potential to detect concealed information. Several previous studies have found that feedback-related brain activities reflect the evaluation process that can be affected by participants' motivation to complete the task. For example, Yeung et al. (2005) showed that the FN was larger in a monetary gambling task when participants had to make a choice than when participants made no active choices, suggesting larger FN was associated with participants' engagement in the task. Recently, Luo et al. (2011) found a larger FN after participants made a deceptive response than when participants made a simple non-deceptive response. This FN may reflect participants' elevated motivation to deceive and to evaluate their performance during the deceptive responses. In the present study, since participants were explicitly instructed to deny the recognition of the concealed personal information, they would expect success feedbacks following the to-be-concealed information as well as success in their deception. This greater motivation in feedback processing thus leads to stronger feedback-related ERPs for feedbacks following probe information than for feedbacks following irrelevant information.

In addition, we found another positivity at right central area that reached maximum amplitude at approximately 400 ms. We found that this positivity was sensitive to feedback valence, with success eliciting a higher positivity than failure. Interestingly, we found this positivity could discriminate probe information from irrelevant information, with feedbacks following a probe eliciting higher positivity than feedbacks following irrelevants. Thus, this positivity encodes both feedback valence (i.e. lie vs. truth) and the stimulus/response that elicits the feedback (i.e. probe vs. irrelevant).

The centroparietal positivity peaking around 453 ms resembles the P300 following FNs (e.g. Foti et al., 2011; Lole et al., 2013). This positivity distinguished between probe-feedbacks and irrelevant-feedbacks, but not feedback valence such as truth or lie. Previous studies suggest that the amplitude of P300 following FN can be affected by participants' motivation, such that the magnitude of reward and participants' involvement in the task led to larger

P300 (Sato et al., 2005; Yeung & Sanfey, 2004; Yeung et al., 2005). Recently, Luo et al. (2011) also found a larger feedback-related P300 after participants made a deceptive response than after participants made a simple response. This centroparietal positivity may represent the similar motivational process. Because participants tried to conceal the probe and avoid being detected by the lie detector, this makes the probe to have a greater motivational significance than irrelevant. This enhanced motivational significance led to larger centroparietal positivity associated with probe.

Based on the present finding, it suggests that the feedback-related ERPs may reflect motivated outcome evaluation that is independent with memory-mediate processes such as item recognition. Indeed, there was no relationship between feedback-related ERPs and recognition-based P300. This pattern of results provides further evidence that the feedback-related ERPs and P300 tap into different psychological processes underlying memory concealment.

In terms of practical implications, CITs that are administered in the field can be benefited from the feedback manipulation. Specifically, there could be more than one way to manipulate real life outcomes depending on specific context. For instance, one can give feedback regarding the probability of being detected or not, e.g., "80% that you will fail the test". One can also use numerical feedback such as +5 or −2, whereas the number can represent years served in prison, or penalty involved if the crime is true. Regardless of the specific feedback manipulation, the idea is to engage in performance evaluation neural processes in addition to memory-related neural process in the CIT.

The present study raises new questions that warrant future investigation. First, previous studies that investigate feedback processing typically use monetary gain or loss as feedbacks (e.g. in a gambling task, Gehring & Willoughby, 2002). In the present study, however, we used "truth" and "lie" as feedback stimuli. Compared to monetary feedbacks that directly contain utilitarian information, "truth" or "lie" resembles more like social feedbacks that one may expect during social interaction. It is possible that our Cz/C4 positivity that encodes truth or lie may reflect such social evaluative processes. Future studies can directly compare monetary feedback and social feedbacks to illustrate the neurocognitive processes underlying these two types of feedback processing. A related question is that although participants were told that they will receive "truth" or "lie" feedbacks, these feedbacks do not have real-life outcomes. Thus, future studies can manipulate the real-life outcome that is linked to such feedbacks (e.g. a reward given a "truth" feedback) to further investigate feedback processing in concealed memory detection. Finally, from an applied view, it will be informative to examine whether combining P300 and feedback-related ERPs can improve discrimination efficiency between guilty and innocent participants in the CIT. The present study provided preliminary evidence that this may be the case: the feedback-related ERPs were not correlated with P300. For innocent participants, given that the probe is simply another irrelevant for innocent participants, there is no reason to expect any systematic differences between the probe and irrelevant feedback-related ERPs. Therefore, combining different ERPs may provide complementary information regarding memory concealment.

5. Conclusion

To conclude, via manipulating feedbacks in a CIT, the present study revealed a series of neural signals during feedback processing that were not only sensitive to feedback valence, but could also identify concealed information. Whereas P300 mainly reflects the recognition and memory–related process, the feedback-related ERPs may reflect the motivational significance of feedback/outcome evaluations during deception or information concealment. These results suggest that the in a feedback CIT context, not only the memory retrieval, but also the motivational processes play an indispensable role in memory concealment.

Acknowledgment

The present study was supported by the National Science Foundation of China to G. Fu (31070894 and 31371041).

References

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

Dien, J. (2012). Applying principal components analysis to event-related potentials: A tutorial. *Developmental Neuropsychology, 37*(6), 497–517. http://dx.doi.org/10.1080/87565641.2012.697503.

Dien, J., Khoe, W., & Mangun, G. R. (2007). Evaluation of PCA and ICA of simulated ERPs: Promax vs. Infomax rotations. *Human Brain Mapping, 28*(8), 742–763. http://dx.doi.org/10.1002/hbm.20304.

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

Donchin, E., & Heffley, E. F. (1979). The independence of the P300 and the CNV reviewed: A reply to Wastell. *Biological Psychology, 9*(3), 177–188. http://dx.doi.org/10.1016/0301-0511(79)90038-3.

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *The Journal of Neuroscience, 32*(35), 12087–12092. http://dx.doi.org/10.1523/JNEUROSCI.1410-12.2012.

Foti, D., Weinberg, A., Dien, J., & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity. *Human Brain Mapping, 32*(12), 2207–2216. http://dx.doi.org/10.1002/hbm.21182.

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

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295*(5563), 2279–2282. http://dx.doi.org/10.1126/science.1066893.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679–708. http://dx.doi.org/10.1037/0033-295X.109.4.679.

Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology, 41*(2), 245–253. http://dx.doi.org/10.1111/j.1469-8986.2004.00152.x.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport, 14*(18), 2481–2484. http://dx.doi.org/10.1097/01.wnr.0000099601.41403.a5.

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45*(5), 688–697. http://dx.doi.org/10.1111/j.1469-8986.2008.00668.x.

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

Hu, X., Pornpattananangkul, N., & Rosenfeld, J. P. (2013). N200 and P300 as orthogonal and integrable indicators of distinct awareness and recognition processes in memory detection. *Psychophysiology, 50*(5), 454–464. http://dx.doi.org/10.1111/psyp.12018.

Hu, X., & Rosenfeld, J. P. (2012). Combining the P300-complex trial-based concealed information test and the reaction time-based autobiographical implicit association test in concealed memory detection. *Psychophysiology, 49*(8), 1090–1100.

Johnson, R. Jr, (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. *Advances in Psychophysiology, 3*, 69–137.

Kubo, K., & Nittono, H. (2009). The role of intention to conceal in the P300-based concealed information test. *Applied Psychophysiology and Biofeedback, 34*(3), 227–235. http://dx.doi.org/10.1007/s10484-009-9089-y.

Lole, L., Gonsalvez, C. J., Barry, R. J., & De Blasio, F. M. (2013). Can event-related potentials serve as neural markers for wins, losses, and near-wins in a gambling task? A principal components analysis. *International Journal of Psychophysiology*. http://dx.doi.org/10.1016/j.ijpsycho.2013.06.011.

Lui, M., & Rosenfeld, J. P. (2009). The application of subliminal priming in lie detection: Scenario for identification of members of a terrorist ring.

Psychophysiology, 46(4), 889–903. http://dx.doi.org/10.1111/j.1469-8986.2009.00810.x.
Luo, Y.-J., Sun, S.-Y., Mai, X.-Q., Gu, R.-L., & Zhang, H.-J. (2011). Outcome evaluation in decision making: ERP studies. In *Culture and neural frames of cognition and communication* (pp. 249–285). Springer. Doi: 10.1007/978-3-642-15423-2_16.
Matsuda, I., Nittono, H., & Ogawa, T. (2013). Identifying concealment-related responses in the concealed information test. *Psychophysiology, 50*(7), 617–626. http://dx.doi.org/10.1111/psyp.12046.
Meijer, E. H., Smulders, F. T., Johnston, J. E., & Merckelbach, H. L. (2007). Combining skin conductance and forced choice in the detection of concealed information. *Psychophysiology, 44*(5), 814–822. http://dx.doi.org/10.1111/j.1469-8986.2007.00543.x.
Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience, 9*(6), 788–798. http://dx.doi.org/10.1162/jocn.1997.9.6.788.
Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience and Biobehavioral Reviews, 28*(4), 441–448. http://dx.doi.org/10.1016/j.neubiorev.2004.05.003.
Rosenfeld, J. P. (2011). P300 in detecting concealed information. *Memory Detection: Theory and application of the Concealed Information Test*, 63–89.
Rosenfeld, J. P., Hu, X., Labkovsky, E., Meixner, J., & Winograd, M. R. (2013). Review of recent studies and issues regarding the P300-based complex trial protocol for detection of concealed information. *International Journal of Psychophysiology, 90*(2), 118–134. http://dx.doi.org/10.1016/j.ijpsycho.2013.08.012.
Rosenfeld, J. P., Hu, X., & Pederson, K. (2012). Deception awareness improves P300-based deception detection in concealed information tests. *International Journal of Psychophysiology, 86*(1), 114–121. http://dx.doi.org/10.1016/j.ijpsycho.2012.06.007.
Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology, 45*(6), 906–919. http://dx.doi.org/10.1111/j.1469-8986.2008.00708.x.
Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., et al. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport, 16*(4), 407–411.
Soskins, M., Rosenfeld, J. P., & Niendam, T. (2001). Peak-to-peak measurement of P300 recorded at 0.3 Hz high pass filter settings in intraindividual diagnosis: complex vs. simple paradigms. *International Journal of Psychophysiology, 40*(2), 173–180.
Spencer, K. M., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology, 38*(2), 343–358. http://dx.doi.org/10.1111/1469-8986.3820343.
Verschuere, B., Rosenfeld, J. P., Winograd, M. R., Labkovsky, E., & Wiersema, R. (2009). The role of deception in P300 memory detection. *Legal and Criminological Psychology, 14*(2), 253–262. http://dx.doi.org/10.1348/135532508X384184.
Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex, 15*(5), 535–544. http://dx.doi.org/10.1093/cercor/bhh153.
Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience, 24*(28), 6258–6264. http://dx.doi.org/10.1523/JNEUROSCI.4537-03.2004.