Neuroscience Letters 467 (2009) 63–66

Contents lists available at ScienceDirect

Neuroscience Letters

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

Does willingness affect the N2-P3 effect of deceptive and honest responses?

Haiyan Wu[a], Xiaoqing Hu[a,b], Genyue Fu[a,*]

[a] Department of Psychology, Zhejiang Normal University, Jinhua 321004, China
[b] Department of Psychology, Northwestern University, Evanston, IL 60208, USA

ARTICLE INFO

Article history:
Received 12 July 2009
Received in revised form
16 September 2009
Accepted 2 October 2009

Keywords:
Willingness
Deception
P300
N2
Conflict detection

ABSTRACT

The present investigation examined the effect of willingness on honest and deceptive responses. Event-related potentials were recorded while participants made deceptive and honest response that were either self-determined or forced. Results showed that the reaction time was faster in response to old words compared to new words and honest responses were faster than deceptive responses. In addition, the P300 of honest responses was significantly more positive than deceptive responses and a significant main effect of willingness indicated that the P300 amplitude, elicited by self-determined responses, was more positive than forced responses. Moreover, the conflict detection N2 component was significantly more negative-going in the lying versus honest responses at Cz. The main effect of willingness also revealed that the forced response evoked a more negative N2 than the self-determined response. These results suggested that deception may involve conflict detection and that there are significant differences in neurological processing between forced deception and self-determined deception.

As an adaptive social behavior, deception is an important part of our daily life. With a recent increased interest in social cognitive neuroscience, a number of investigators have started to explore the underlying neural mechanisms of deception [18,17,1,8]. However, one general limitation of these recent paradigms, as stated by Sip et al. [15], is that participants were directed to make deceptive response in laboratory settings. Although this "directed" deception constitutes some core elements of deception (e.g., response inhibition [10,9,11]), the true complexity naturally produced lies remains unclear.

The present investigation attempts to address this limitation by examining the influence of willingness on participants' deceptive behavior. In this modified paradigm we asked participants to respond either deceptively or honestly (deception vs. truth) to two successive words in a trial. Participants were to respond by indicating whether the word was one they had seen previously (old) or a new word (new). The first response was always self-determined, while the second response was to be the opposite response of the first. For example, if a participant responded deceptively in response to the first word, they should be honest in response to the second word. Since participants were given no choice for their response to the second word it was considered a forced response.

The present study also examined the possible N2-P3 effect when making deceptive responses. Previous Event Related Potential (ERP) deception detection studies have repeatedly demonstrated that concealing information elicits a considerable P300 and that this component is a good indicator of deception [2,16,13,12,14]. The rationale behind the P300-deception detection is that the memorized information will stand out in a series of un-memorized stimuli. Previous studies have found support for this P300 in old/new effect. Thus, in the present study, we expect to obtain the old/new effect by manipulating word type (memorized vs. new words).

Previous neuroimaging research has found executive functioning, as reflected by Anterior Cingulate Cortex (ACC) and lateral prefrontal activation, as well as Medial Frontal Negativities (MFN) enhancement, is a key component in deception. However, the frontal–central N2 component, which is regarded as reflecting conflict detection [7,5,6,3], has not yet been reported in ERP deception studies. As a function of response mode (deception vs. truth), we hypothesized that the N2 effect would emerge when making deception responses. Moreover, it remains unclear whether there is a possible difference between self-determined deception and forced deception.

To summarize, the goal of the current study was to manipulate the variable of willingness to deceive (self-determined vs. forced) and investigate the N2-P3 effect of self-determined and forced deception.

15 university students (M = 21.18 years, SD = .77; range: 12–22; 6 males) of Zhejiang Normal University in the P.R. China took part in the experiment. All participants were native Chinese speakers, with normal or corrected-to-normal vision, non-regular substance users and currently not under the influence of drugs, alcohol, or

* Corresponding author: Genyue Fu, Department of Psychology, Zhejiang Normal University, Jinhua, Zhejiang, China.
E-mail address: fugy@zjnu.cn (G. Fu).

0304-3940/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.neulet.2009.10.002

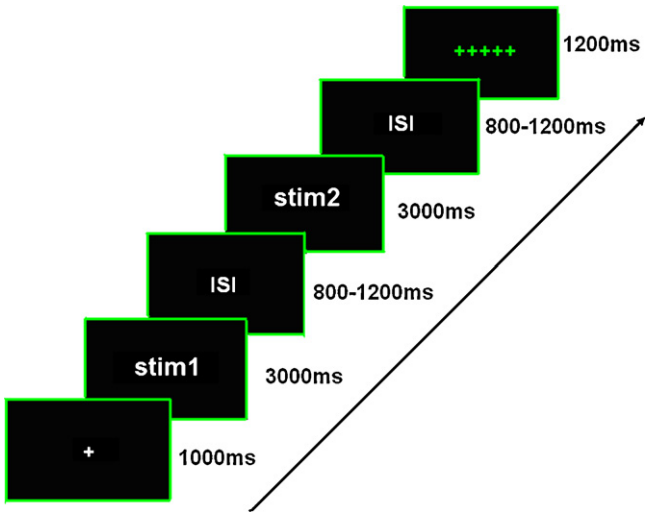

Fig. 1. The experimental procedure of one trial in present study.

psychotropic medications, and free of psychological disorders or disorders known to affect the CNS (e.g., previous head injury resulting in the loss of consciousness). All participants provided informed consent prior to the experiment and received ¥ 10 per hour for participation.

130 words were selected for stimuli in this study. They consisted of 65 "old" items (i.e., direct subjects to remember) and 65 "new" words (i.e., subjects were not asked to view prior to the experiment). All the words were matched in frequency, emotion value and semantic category.

Prior to the experimental session, participants were given a list of 65 unrelated words to memorize ("old" words). Then a word recognition test was given to ensure participants memorized the words. Participants did not proceed to the next step until their word recognition accuracy was above 95%.

Five new words and five old words were selected from the new and old words list for the practice phase and were never presented in formal experiment. Consequently, there are 60 new and 60 old words in experimental phase. All new and old words were presented 1:1 in three blocks and the order of the first and second word was counterbalanced. Four different word-pair categories were created (new–new, old–old, new–old, old–new). Each of the three blocks included 80 word trials. In each block the ratio of the four word-pair categories was equal.

Participants arrived at the lab on the day of the experiment. They memorized 65 words and were tested in an old/new recognition task that included all stimuli used in the study.

Following the word recognition test, subjects were taken to the dimly lit room and conducted the practice session with five trials of deceptive and honest responses. Then the experimental phase began with ERP recording. Participants were given the following instructions: "Please sit still and pay attention to the stimuli and judge whether the word is new or old. But please note that: there are two stimuli in a trial, you should choose by yourself to answer deceptively or honestly to the first word and then conduct the opposite response to the second word" (see Fig. 1). Participants were also told to ensure the ratio of deception to the first and the second word is 1:1. Accordingly, the self-determined response is defined as either honest or deceptive response to the first word because participants made the choice themselves and the forced response is the response to the second word. Participants were asked to blink only in allotted time window.

Stimuli were presented with E-prime software and displayed in the center of a monitor at a viewing distance of 60 cm. All words were written in white on a black background. The inter-stimulus interval (ISI) ranged from 800 to 1200 ms randomly.

Electroencephalogram (EEG) activity was recorded from scalp electrodes using a 128-channel Geodesic Sensor Net and Energy and Geoscience Institute (EGI) software. Electrode impedances were kept below 50 kΩ before recording. All recordings were referenced to Cz, and an averaged reference was calculated offline. Signals were sampled at 500 Hz and EEG data were filtered using a 0.1–30 Hz bandpass offline. All incorrect trials (i.e., make truthful or deceptive response to both words in a single trial) were excluded from further analysis. Continuous EEG was segmented into four condition-related stimulus-lock epochs from 100 ms before to 1000 ms after stimuli: (1) self-determined honesty with truthful response; (2) self-determined deception with deceptive response; (3) forced honesty with truthful response; (4) forced deception with deceptive response.

We excluded trials with blinking and eye movement artifacts and trials in which 20 or more channels exceeded a voltage threshold of 100 μV (absolute) or a transition threshold of 100 μV (sample to sample). The total artifact-free percent is 88.04%. Correct, artifact-free trials were averaged for each subject in each condition, and the data were baseline-corrected 100 ms before stimulus onset. SPSS 11.0 package is used for statistical analysis and Analysis of Variance (ANOVA) results were corrected using the Greenhouse-Geisser procedure whenever the sphericity assumption was violated. In total, 3 subjects were excluded because of more than 100 bad segmentations, so the results presented here were based on the remaining 12 participants.

Table 1 contains the behavioral responses of eight different categories: self-determined deception to old items, self-determined deception to new items, self-determined honesty to old items, self-determined honesty to new items, forced deception to old items, forced deception to new items, forced honesty to old items, and forced honesty to new items. A 2 willingness (self-determined, forced) × 2 response (honest, deceptive) × 2 word type (old, new) ANOVA on RT revealed a significant main effect of response, $F(1,11) = 18.491$, $p = 0.001$, $\eta_p^2 = 0.627$, with participants responding significantly slower during deception ($M = 1380$ ms) compared to honest responses ($M = 1271$ ms). In addition, a main effect of word type was found, $F(1,11) = 38.672$, $p < 0.001$, $\eta_p^2 = 0.779$, indicating a faster reaction time for old words ($M = 1254$ ms) than to new words ($M = 1397$ ms). No other significant main effects or interactions were found.

Fig. 2 shows the grand averaged ERP waveforms at Fz and Cz. As shown in Fig. 2, frontal–central ERP activity with a N2-P3 distribution.

With respect to the N2, a frontal–central negative component reflecting the conflict detection by the ACC emerges around 250–300 ms post-stimulus. Considering the longer RT time in our study, a 250–400 ms post-stimulus time window was selected for the statistic. A 2 electrode site (Fz, Cz) × 2 willingness (self-determined, forced) × 2 response (deceptive,

Table 1
Mean (SD) RT for both new and old words of each condition.

	Self-determined deception	Self-determined honesty	Forced deception	Forced honesty
Old	1367.3 (459.3)	1151.4 (390)	1310.2 (344.1)	1186.9 (227.9)
New	1406.1 (481)	1408.5 (497)	1437.3 (414.3)	1337.3 (286.1)

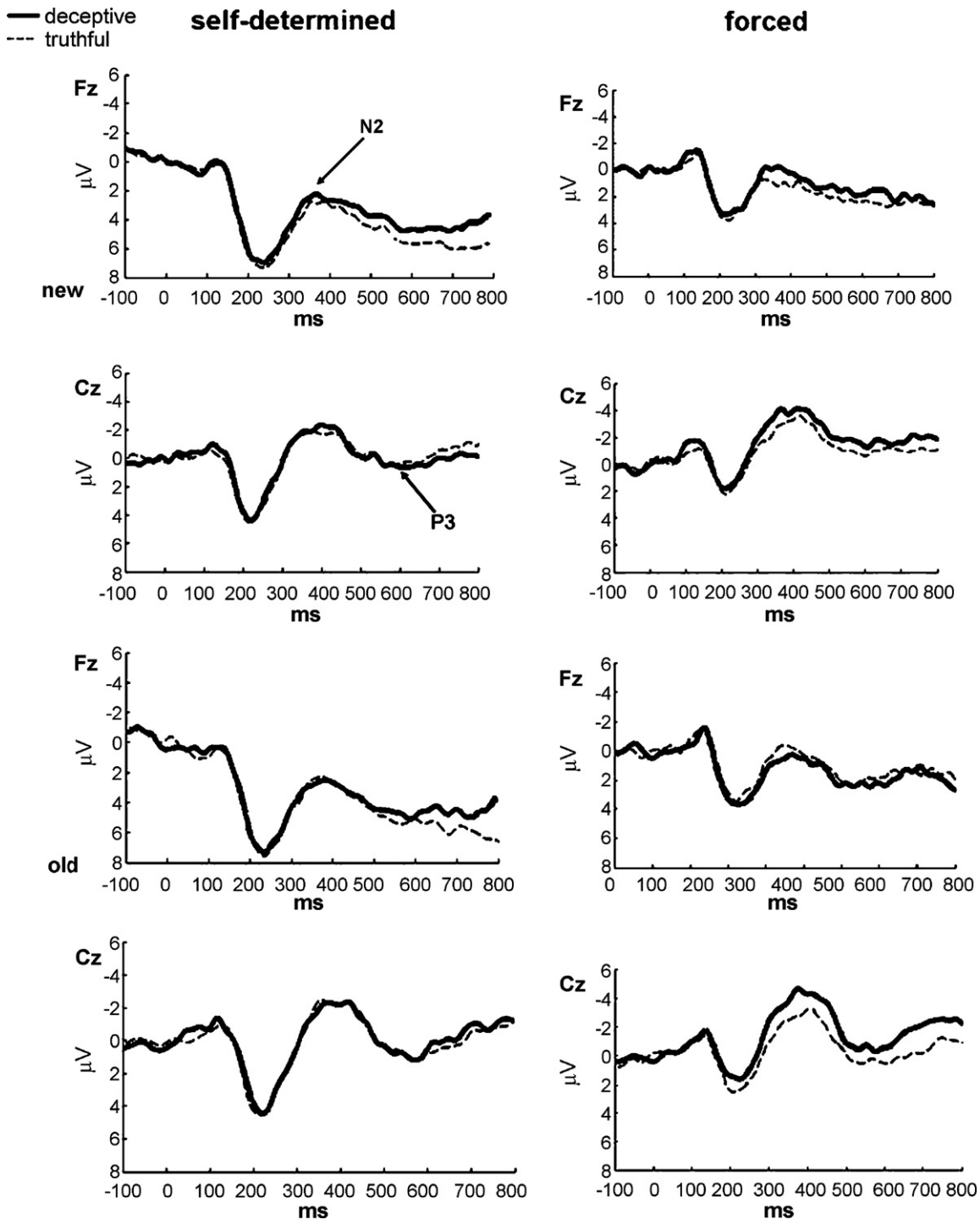

Fig. 2. Grand averaged waveforms at Fz and Cz.

true) × 2 word type (new, old) repeated-measures ANOVA on N2 amplitude revealed a significant main effect of electrode sites, $F(1,11) = 29.326$, $p < 0.001$, $\eta_p^2 = 0.727$, indicating a negative-going deflection that was significantly more negative at Cz ($M = -3.223\ \mu V$) than Fz ($M = 0.841\ \mu V$). A main effect of willingness was also found, $F(1,11) = 11.537$, $p < 0.02$, $\eta_p^2 = 0.512$, indicating that forced responses ($M = -1.898\ \mu V$) elicited a significantly more negative N2 than self-determined responses ($M = 1.898\ \mu V$). A significant main effect of response, $F(1,11) = 5.959$, $p < 0.05$, $\eta_p^2 = 0.351$, demonstrated that deceptive responses ($M = -1.367\ \mu V$) evoked a more negative N2 than honest responses ($M = -1.015\ \mu V$). The main effects of electrode sites and response were qualified by a significant interaction, $F(1,11) = 3.549$, $p < 0.05$, $\eta_p^2 = 0.324$. Follow-up analyses revealed that the N2 of deceptive response ($M = -3.534\ \mu V$) was significantly more negative than honest response ($M = -2.911\ \mu V$) at Cz, $F(1,11) = 19.396$, $p = 0.001$, $\eta_p^2 = 0.638$, while there was no significant difference at Fz.

P300 is also known as the Late Positive Component with the peak amplitude from 400 ms to 800 ms after stimulus onset. Building on the N2 effect, we select the 400–800 ms time window for P300 at Fz and Cz. A 2 electrode site (Fz, Cz) × 2 willingness (self-determined, forced) × 2 response (deceptive, true) × 2 word

type (new, old) repeated-measures ANOVA on the P300 amplitude revealed significant main effects of electrode sites [$F(1,11) = 12.594$, $p < 0.01$, $\eta_p^2 = 0.534$], willingness [$F(1,11) = 12.412$, $p < 0.01$, $\eta_p^2 = 0.54$], response [$F(1,11) = 22.594$, $p < 0.005$, $\eta_p^2 = 0.673$] and word type [$F(1,11) = 6.472$, $p < 0.05$, $\eta_p^2 = 0.370$]. The electrode sites effects showed that the P300 amplitude over Fz ($M = -0.702\,\mu V$) was larger than Cz ($M = -4.66\,\mu V$). The significant main effect of willingness indicating than the P300 amplitude elicited by a self-determined response ($M = 3.031\,\mu V$) was more positive than a forced response ($M = 2.331\,\mu V$). The main effect of response showed that an honest response ($M = 3.622\,\mu V$) was more positive than a deceptive response ($M = 1.74\,\mu V$). Moreover, old words evoked a P300 amplitude ($M = 2.895\,\mu V$) that was more positive than new words ($M = 2.467\,\mu V$). No significant interactions were found.

The behavior results showing that overall deceptive responses were associated with slower RTs than honest responses confirmed the results of Johnson et al. [9]. In addition, the EEG results demonstrating that deception involved larger N2s (at Cz) and smaller P300 (at Fz and Cz) was consistent with the general conclusion that deception requires both response inhibition and a greater workload.

Our findings prove support for the validity of the P300 amplitude to distinguish between deception and the truth regardless of stimulus type or willingness. Consistent with previous studies [12,19], the deception-related P300 was less positive than honest response over frontal–central sites and the P300 of old words was significantly more positive. Our findings also suggest self-determined responses elicited more positive P300 than forced response, especially to new words. Two factors could contribute to these differences in P300s. First, self-determined responses were made in response to the first words, which may have been more novel than the second words (forced response). This novelty could have resulted in a more positive P300. Another possible explanation is that, upon the presentation of the second words (forced response), participants are required to respond in a specific manner, this may have placed a higher cognitive workload on the response inhibition process compared to the first stimulus.

As an index of participants' conflict detection, the N2 elicited by deceptive response was found to be more negative than the honest response. In addition, the forced response was more negative than self-determined response, possibly due to an increased response inhibition process. These results are closely related to the P300 results and can also be explained by the response conflict associate with an increased workload. These results are also consistent with prior work demonstrating that deceptive response need more executive control process which can be reflected in N2 through conflict inhibition [7,5,4]. Combined with the P300 results, the more negative N2 amplitude pattern of the forced response may involve more conflict control than self-generate responses. The results showed that forced responses are linked to larger N2 and smaller P3 which may also indicate that the ecological validity of the lie detection studies using "directed lies" could be harmed in field use where the lies are self-determined. However, the present study did not find a willingness interaction with the response. Further research is needed to clarify whether there is a possible mechanism for different types of deception and possible interaction effects between response type and willingness to deceive.

In conclusion, the present findings demonstrated that the P300 reflected the old/new effect and was related to mental workload. In addition, deceptive responses are associated with more negative N2 (central) and less positive P300 (frontal–central), which was independent of the effect of willingness. Moreover, our investigation revealed that forced deception and forced honesty are related to a more negative N2 and less positive P300 than self-determined deception and honesty. Together, these behavioral and ERP findings are consistent with the concept that making deceptive statements increases response conflict and involves a higher cognitive workload, especially for forced response.

Acknowledgements

This work was supported by The National Natural Science Foundation of China (No. 30670696), Program for Innovative Research Team in Zheijiang Normal University and Innovation Project for Graduate students of Zhejiang Province, China. The author would like to thank Dr. Angela Evans for her suggestions on the manuscript of this paper.

References

[1] N. Abe, M. Suzuki, T. Tsukiura, E. Mori, K. Yamaguchi, M. Itoh, T. Fujii, Dissociable roles of prefrontal and anterior cingulate cortices in deception, Cerebral Cortex 16 (2006) 192–199.
[2] L.M. Ann, J.P. Rosenfeld, A.H. Ryan, Thirty-site P300 scalp distribution, amplitude variance across sites, and amplitude in detection of deceptive concealment of multiple guilty items, Social Neuroscience 1 (2008) 1–19.
[3] A. Dimoska, S.J. Johnstone, R.J. Barry, The auditory-evoked N2 and P3 components in the stop-signal task: indices of inhibition, response-conflict or error-detection? Brain and Cognition 62 (2006) 98–112.
[4] G. Dong, L. Yang, Y. Hu, Y. Jiang, Is N2 associated with successful suppression of behavior responses in impulse control processes? Neuroreport 20 (2009) 537–542.
[5] F.C.L. Donkers, van G.J.M. Boxtel, The N2 in go/no-go tasks reflects conflict monitoring not response inhibition, Brain and Cognition 56 (2004) 165–176.
[6] M. Falkenstein, Inhibition, conflict and the Nogo-N2, Clinical Neurophysiology 117 (2006) 1638–1640.
[7] J.R. Folstein, P.C. Van, Influence of cognitive control and mismatch on the N2 component of the ERP: a review, Psychophysiology 45 (2008) 152–170.
[8] G. Ganis, S.M. Kosslyn, S. Stose, W.L. Thompson, D.A. Yurgelun-Todd, Neural correlates of different types of deception: an fMRI investigation, Cerebral Cortex 13 (2006) 830–836.
[9] R.J. Johnson, J. Barnhardt, J. Zhu, The deceptive response: effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity, Biological Psychology 64 (2003) 217–253.
[10] R.J. Johnson, J. Barnhardt, J. Zhu, The contribution of executive processes to deceptive responding, Neuropsychologia 42 (2004) 878–901.
[11] R.J. Johnson, H. Henkell, E. Simon, J. Zhu, The self in conflict: the role of executive processes during truthful and deceptive responses about attitudes, Neuroimage 39 (2008) 469–482.
[12] A.R. Miller, J.P. Rosenfeld, M. Soskins, M. Jhee, P300 amplitude and topography in an autobiographical oddball paradigm involving simulated amnesia, Psychophysiology 16 (2002) 1–11.
[13] J.P. Rosenfeld, A.M. Reinhart, M. Bhatt, J. Ellwanger, K. Gora, M. Sekera, J. Sweet, P300 correlates of simulated malingered amnesia in a matching-to-sample task: topographic analyses of deception versus truthtelling responses, International Journal of Psychophysiology 28 (1998) 233–247.
[14] J.P. Rosenfeld, J.R. Biroschak, J.J. Furedy, P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms, International Journal of Psychophysiology 60 (2006) 251–259.
[15] K.E. Sip, A. Roepstorff, W. McGregor, C.D. Frith, Detecting deception: the scope and limits, Trends in Cognitive Sciences 12 (2008) 48–53.
[16] M. Soskins, J.P. Rosenfeld, T. Niendam, 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 (2001) 173–180.
[17] S.A. Spence, T.F. Farrow, A.E. Herford, I.D. Wilkinson, Y. Zheng, P.W. Woodruff, Behavioural and functional anatomical correlates of deception in humans, Neuroreport 12 (2001) 2849–2853.
[18] S.A. Spence, M.D. Hunter, T.F. Farrow, R.D. Green, D.H. Leung, C.J. Hughes, V. Ganesan, A cognitive neurobiological account of deception: evidence from functional neuroimaging, Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 359 (2004) 1755–1762.
[19] H.P. Tardif, R.J. Barry, A.M. Fox, S.J. Johnstone, Detection of feigned recognition memory impairment using the old/new effect of the event-related potential, International Journal of Psychophysiology 36 (2000) 1–9.