Acta Biologica Hungarica 65(1), pp. 1–12 (2014)
DOI: 10.1556/ABiol.65.2014.1.1

AUTOMATIC DETECTION OF TRUSTWORTHINESS OF THE FACE: A VISUAL MISMATCH NEGATIVITY STUDY

Z. Kovács-Bálint,[1] G. Stefanics,[2,3] A. Trunk[1] and I. Hernádi[1,*]

[1]Department of Experimental Zoology and Neurobiology, Institute of Biology, University of Pécs, Hungary;
[2]Translational Neuromodeling Unit, Institute for Biomedical Engineering, University of Zurich & ETH Zurich, Switzerland;
[3]Laboratory for Social and Neural Systems Research, Institute for Empirical Research in Economics, University of Zürich, Switzerland

(Received: June 12, 2013; accepted: June 24, 2013)

Recognizing intentions of strangers from facial cues is crucial in everyday social interactions. Recent studies demonstrated enhanced event related potential (ERP) responses to untrustworthy compared to trustworthy faces. The aim of the present study was to investigate the electrophysiological correlates of automatic processing of trustworthiness cues in a visual oddball paradigm in two consecutive experimental blocks. In one block, frequent trustworthy ($p = 0.9$) and rare untrustworthy face stimuli ($p = 0.1$) were briefly presented on a computer screen with each stimulus consisting of four peripherally positioned faces. In the other block stimuli were presented with reversed probabilities enabling the comparison of ERPs evoked by physically identical deviant and standard stimuli. To avoid attentional effects participants engaged in a central detection task. Analyses of deviant minus standard difference waveforms revealed that deviant untrustworthy but not trustworthy faces elicited the visual mismatch negativity (vMMN) component. The present results indicate that adaptation occurred to repeated unattended trustworthy (but not untrustworthy) faces, i.e., an automatic expectation was elicited towards trustworthiness signals, which was violated by deviant untrustworthy faces. As an evolutionary adaptive mechanism, the observed fast detection of trustworthiness-related social facial cues may serve as the basis of conscious recognition of reliable partners.

Keywords: Trustworthiness – social perception – event-related potential (ERP) – visual mismatch negativity (vMMN) – preattentive processing

INTRODUCTION

In everyday life people infer personality traits from one another's behavior. Especially, facial expressions are recognized as key aspects of social interactions [4, 5]. Recent investigations about general trustworthiness revealed that specific brain areas may serve the evolutionary adaptive cognitive mechanism to recognize untrustworthy counterparts. [8, 23]. It is also known that a very short exposure (approximately 100–200 ms) is already sufficient to properly evaluate a face along its trustworthiness dimension [29, 35, 36].

*Corresponding author; e-mail address: hernadi@gamma.ttk.pte.hu

Prior studies have concluded that the encoding of social facial features may occur at a fairly early stage of processing, perhaps in parallel with structural encoding of facial features [11, 37]. It has been accepted that facial expressions related to trustworthiness modify face-sensitive event related potential (ERP) components. Recent studies revealed that untrustworthy faces elicit higher P100 response (at approx. 60–100 ms after stimulus onset) compared to trustworthy faces [23, 37].

Several studies using the visual mismatch negativity (vMMN) paradigm have shown that the vMMN component is a useful index to study automatic pre-attentive change detection processes in the brain (for a recent review see: [17]). As being a counterpart of the auditory mismatch negativity [24], the vMMN response can be typically elicited by infrequent (deviant) stimuli presented among frequent (standard) ones, e.g., by deviant color [9], orientation [1], movement [26], spatial frequency [15], contrast [30] and even by abstract sequential irregularities [31], and independently of attention [3]. The sources of vMMN generators have been identified in a set of areas comprising occipital, parietal and frontal brain regions [7, 18, 32]. Recently, the rapid processing of emotional expressions of unattended faces was also evidenced by using the vMMN paradigm [2, 14, 18, 33, 39].

Although the above studies have demonstrated that the human brain automatically detects changes in basic emotional expressions without conscious attention, little is known about the automatic processing of trustworthiness cues on the face, which may be relevant to recognizing various key intentions accompanying social interactions. In the present study, we aimed at investigating the time course, automaticity and potentially differential processing of facial cues of trustworthiness and untrustworthiness. First, we hypothesized if trustworthiness cues are processed automatically, i.e. outside of the attentional focus, then frequently repeated presentation of trustworthy faces (standard stimulus) will establish an automatic representation of regularity which will be violated by occasional untrustworthy faces (deviant stimulus). We also expected deviants to elicit the vMMN component as it is considered to be a reliable index of early pre-attentive change detection in the brain. Furthermore, in a reverse block, the probability of untrustworthy and trustworthy faces were swapped, and the reversed hypothesis whether frequently repeated untrustworthiness (standard stimulus) would establish an automatic representation of regularity was also tested.

MATERIALS AND METHODS

Participants

Twenty healthy human volunteers were recruited for a single session ERP study. All participants were right handed, and all of them had normal or corrected-to-normal vision. Data from 5 participants were excluded from the final analysis, either due to low performance in the central task** (detection rate under 85%, n = 3), or due to technical problems with the recordings (n = 2). The final data sample comprised of 15 participants (10 females; mean age 21.27 ± 0.91 yrs). The protocol of the study was approved by the Ethical Committee of the University of Pécs and conformed to the Declaration of Helsinki on human experiments.

Stimuli and procedure

Forty computer-generated faces of different identities were used to create visual stimuli. Faces were selected from the Trustworthiness Face Database (previously developed with FaceGen Modeller 3.1 by Oosterhof and Todorov [25]). The database contained standardized trustworthy and untrustworthy faces, which was developed to reliably measure trustworthiness based on trait cues on the face. For the present study, 40 faces (20 moderate trustworthy and 20 moderate untrustworthy) were chosen from the database. We avoided selecting extreme facial expressions, i.e., those with judgment scores higher than ±2 SD compared to the mean of their category (for details, see [25]). A pixel-based image analysis was additionally performed to ensure that the selected faces did not differ in size, color or luminance. Each stimulus consisted of faces of four different individuals (2 males and 2 females) belonging to the same category (either untrustworthy or trustworthy). Faces were presented in the periphery, in the four visual quadrants of a computer screen with a viewing angle of approx. 6° by 4°, on a uniform gray background. The presentation order of the stimuli was randomized with the restriction that the same faces were not presented on two consecutive stimuli. Stimuli were presented for 150 ms following an inter stimulus interval (ISI) of 500–550 ms. The duration of stimulus presentation and ISI were adopted from a previous study [33], where robust vMMN was observed to deviant facial stimuli (Fig. 1). A total of 2000 stimuli were presented in two consecutive experimental blocks. In one block, standard stimuli (P = 0.9) consisted of trustworthy faces, and deviant stimuli (P = 0.1) consisted of untrustworthy faces (block of rare untrustworthy faces: *Block RUF*). In the other block, the probabilities of the stimulus categories were reversed (block of rare trustworthy faces: *Block RTF*). The presentation order of the two experimental blocks was counterbalanced across participants. In the center of each stimulus screen, a white fixation cross was presented with one line longer than the other. The cross was occasionally flipped by 90 degrees, and the participants' task was to detect the cross-flips which occurred in 10% of the ISIs, and respond with pressing a button with their right hand, while ignoring the faces presented independently in the background (Fig. 1). Reaction time (RT) to cross-flips was recorded as an objective measure of overt attention to the central detection task.

**Relatively low detection rate might indicate that participants did not attend the central fixation cross, therefore to avoid potential attentional confounds in ERPs to peripherally presented stimuli, which might have been attended to during some of the trials, participants with detection rate under 85% were excluded.

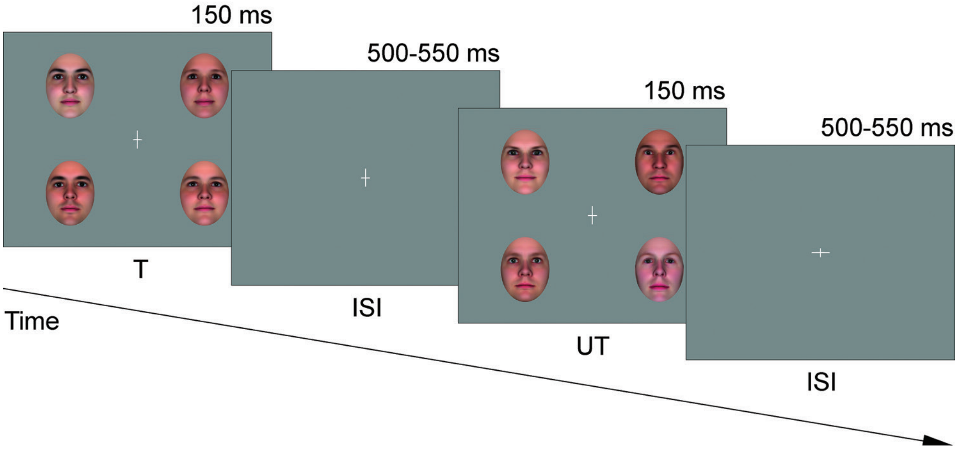

Fig. 1. Block design of the experiment with two sample stimulus panels. Each stimulus consisted of four images of either trustworthy (T) or untrustworthy (UT) faces (two males and two females in each stimulus, in interchanging positions). Stimuli were presented for 150 ms following an inter-trial interval (ISI, 500-550 ms). The participants' task was to indicate a central cross-flip (in 10% of the ISIs, indicated in the right ISI panel)

EEG recording

EEG was recorded from 13 scalp locations (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, O2, T3, T4) with Ag/AgCl electrodes mounted in an electrode cap according to the international 10/20 standard. EOG was recorded from two additional electrodes placed above and below the left and right outer canthi, respectively. Impedances were kept below 4 kΩ. All electrodes were referenced to the nose, and the forehead served as ground. Recording was continuous with an analog band-pass from 0.16 Hz to 150 Hz. EEG was digitized at 1 kHz sampling rate with a 16-bit precision (Power1401, CED, Cambridge, UK) and stored on PC. Data were filtered off-line between 0.5–30 Hz before ERP analysis.

Data analysis

Hit rate and RT for cross-flips were averaged for each participant for each block, and were analyzed by repeated measures analyses of variances (rANOVAs) to compare RT or hit rates between blocks (BLOCK [rare untrustworthy vs. rare trustworthy]) on the basis of the presentation order of the blocks (PRESENTATION ORDER [first block vs. second block]).

EEG was analyzed in Matlab (MathWorks, Natick, MA) using the EEGLAB toolbox [10]. Epochs of –100 to 500 ms were extracted and baseline-corrected for the –100 to 0 time range. Epochs were averaged off-line, separately for standard and deviant stimuli for each block. To avoid any potential confounding effects arising from motor activity related to button pressing, epochs immediately preceding or following responses to cross-flips were excluded from further analysis.

Difference waves were calculated by subtracting ERPs elicited by standard stimuli from ERPs elicited by deviant stimuli. Importantly, ERPs evoked by physically identical stimuli were compared, e.g., differential activity for trustworthy faces was calculated by subtracting ERPs to standard trustworthy faces from ERPs to deviant trustworthy faces. In agreement with previous studies [16, 31], vMMN amplitude measurements were based on grand mean vMMN curves. Thus, vMMN amplitude measurements were done in 20 ms time windows around the two previously identified negative peaks, by calculating the mean amplitude between 115 and 135 ms and 225 and 245 ms, respectively. To enable analysis of lateralized effects, average amplitudes in both selected time windows were calculated for six electrode locations: F3, F4, C3, C4, O1 and O2, respectively. Occipital electrode sites were analyzed as the typical vMMN component appears over these regions; and frontal and central electrodes were analyzed due to the possible involvement of anterior cortical areas in processing of social aspects of face information. First, ERPs were subjected to rANOVA of CONDITION (standard vs. deviant) × ANTERIORITY (anterior vs. central vs. posterior electrodes) × LATERALITY (left vs. right hemisphere), separately for the two stimulus categories (untrustworthy and trustworthy). To test for any condition-specific processing of standard and deviant faces, ERPs were compared by rANOVA of CATEGORY (trustworthy vs. untrustworthy) × ANTERIORITY (anterior vs. central vs. posterior electrodes) × LATERALITY (left vs. right hemisphere), separately for the two conditions (standard and deviant). Where interactions between variables were statistically significant, post hoc Tukey HSD tests were performed to determine pairwise differences. The threshold for significance was set at $\alpha = 0.05$. Greenhouse-Geisser correction of the degrees of freedom was applied where appropriate and corresponding ε values are reported.

RESULTS

Behavioral results

Statistical analysis on average cross flip hit rates (*Block RTF*: 95.4 ± 0.90 per cent vs. *Block RUF*: 94.4 ± 0.87 per cent), or on RT data (*Block RTF*: 384.69 ± 4.62 ms vs. *Block RUF*: 386.54 ± 5.16 ms) did not show significant differences between blocks. Results of behavioral data analysis indicated that different experimental conditions did not have differential effects on the detection accuracy of the cross-flips. High hit rates for the cross-flips suggested that participants allocated great amount of attention to the distracter task; and that the observed brain responses to the difference between deviant and standard face stimuli cannot be accounted for by potential attentional effects.

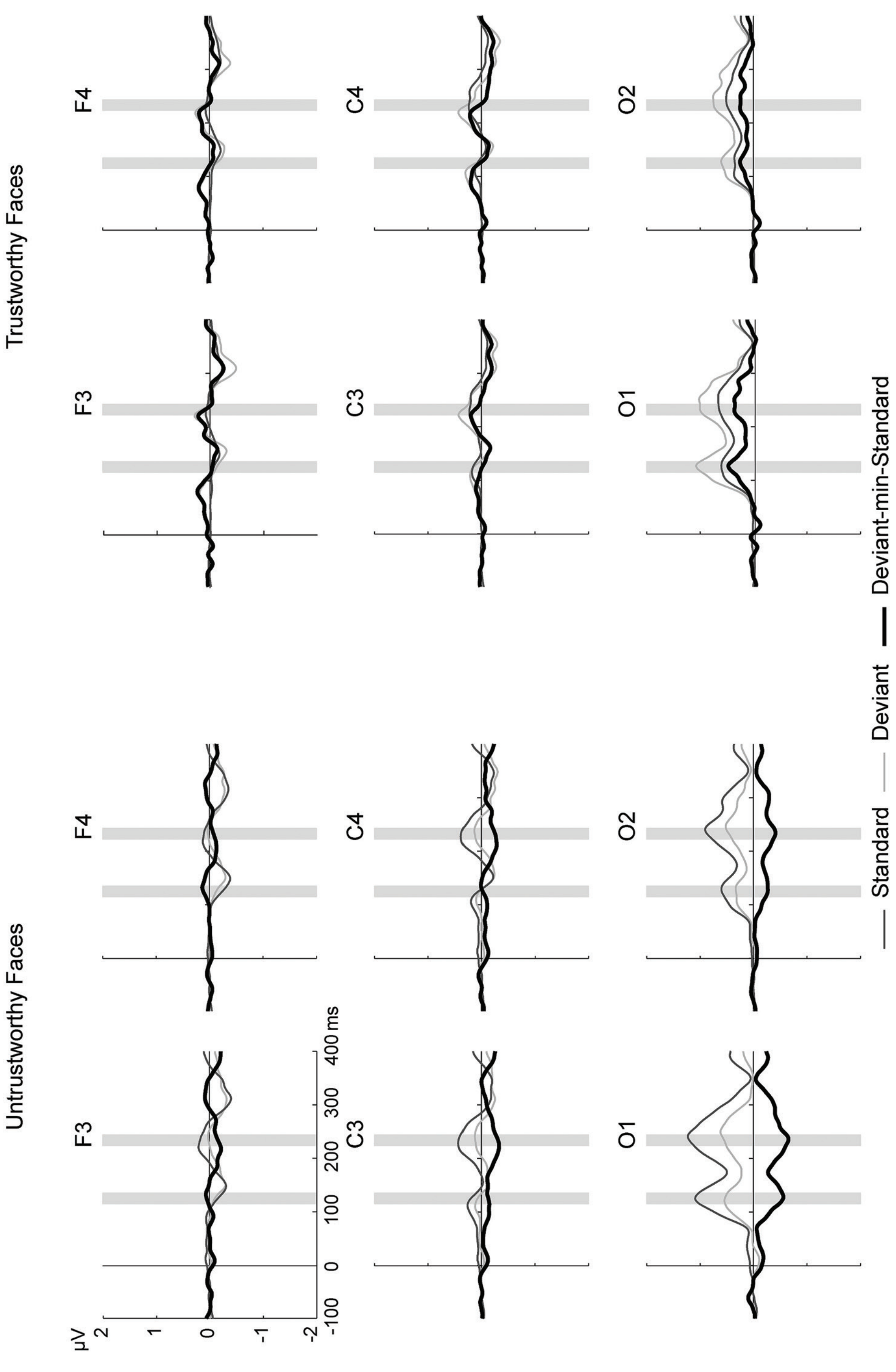

Fig. 2. Grand average (n = 15) event-related potential (ERP) responses and vMMN difference potentials evoked by untrustworthy (left) and trustworthy (right) stimulus categories. Shaded areas mark the statistically analyzed amplitude intervals (115–135 ms and 225–245 ms)

Electrophysiological results

Figure 2 shows grand average ERPs, where shaded areas indicate time windows used for amplitude measurements.

In the earlier (115–135 ms) time window, ANOVA of the amplitude values for ERPs evoked by standard and deviant untrustworthy faces yielded a significant main effect of ANTERIORITY ($F(2,28) = 17.26$, $p < .01$, $\eta^2 = .55$), caused by higher average amplitudes over the occipital electrode sites, and a main effect of LATERALITY ($F(1,14) = 6.63$, $p < .05$, $\eta^2 = .32$), due to higher average amplitudes over the left hemisphere. A significant CONDITION × LATERALITY interaction ($F(1,14) = 4.43$, $p = .054$, $\eta^2 = .24$) was also observed. Post hoc test revealed that the interaction was caused by more negative ERP amplitudes to deviant compared to standard untrustworthy faces over the left hemisphere (Tukey HSD: $p < .01$, Fig. 3A). A significant CONDITION × ANTERIORITY interaction was also observed ($F(2,28) = 3.24$, $p = .054$, $\eta^2 = .19$). Post hoc test revealed a tendency of more negative responses elicited by deviant compared to standard untrustworthy stimuli over posterior electrode sites (Tukey HSD: $p = .06$, Fig. 3B). No other effects or interactions reached significance.

Analysis of amplitude values in the same time window (115–135 ms) evoked by standard and deviant trustworthy faces yielded a significant main effect of ANTERIORITY ($F(2,28) = 18.38$, $p < .01$, $\eta^2 = .57$) and LATERALITY ($F(1,14) = 8.39$, $p < .05$, $\eta^2 = .37$), caused by higher average amplitude values over posterior electrode sites and over the left hemisphere, respectively. No other effects or interactions reached significance.

Comparison of the standard trustworthy and untrustworthy ERP waves in the same (115–135 ms) time window revealed no significant main effect of CATEGORY.

Analysis of the deviant trustworthy and untrustworthy ERPs in the same time window (115–135 ms) yielded significant three-way CATEGORY × LATERALITY × ANTERIORITY interaction ($F(2,28) = 3.58$, $p < .05$, $\eta^2 = .2$). Post hoc test revealed difference between deviant trustworthy and untrustworthy ERPs over O1 and O2 electrode locations (Tukey HSD test: $p<.01$), which was caused by the more negative responses to untrustworthy faces compared to trustworthy faces.

In the later (225–245 ms) time window, ANOVA of the ERP amplitudes for standard and deviant untrustworthy faces yielded marginally significant main effect of CONDITION ($F(1,14) = 3.81$, $p = .07$, $\eta^2 = .21$), caused by less positive amplitude values elicited by deviant compared to standard faces. A significant main effect of ANTERIORITY ($F(2,28) = 7.08$, $p < .01$, $\eta^2 = .34$) revealed that ERP amplitudes were higher over posterior electrode sites compared to the frontal ones; and a significant main effect of LATERALITY ($F(1,14) = 4.95$, $p < .05$, $\eta^2 = .26$) showed that ERP amplitudes were larger over the left hemisphere compared to the right. A marginally significant interaction of CONDITION × LATERALITY was also observed ($F(1,14) = 4.28$, $p = .057$, $\eta^2 = .23$), due to smaller amplitudes elicited by standard stimuli over the right hemisphere (Tukey HSD test: $p < .01$, Fig. 3A). Thus, the left vs. right differences of the ERP responses to the standard (and not to deviant) stimuli may be accounted for the observed vMMN difference wave.

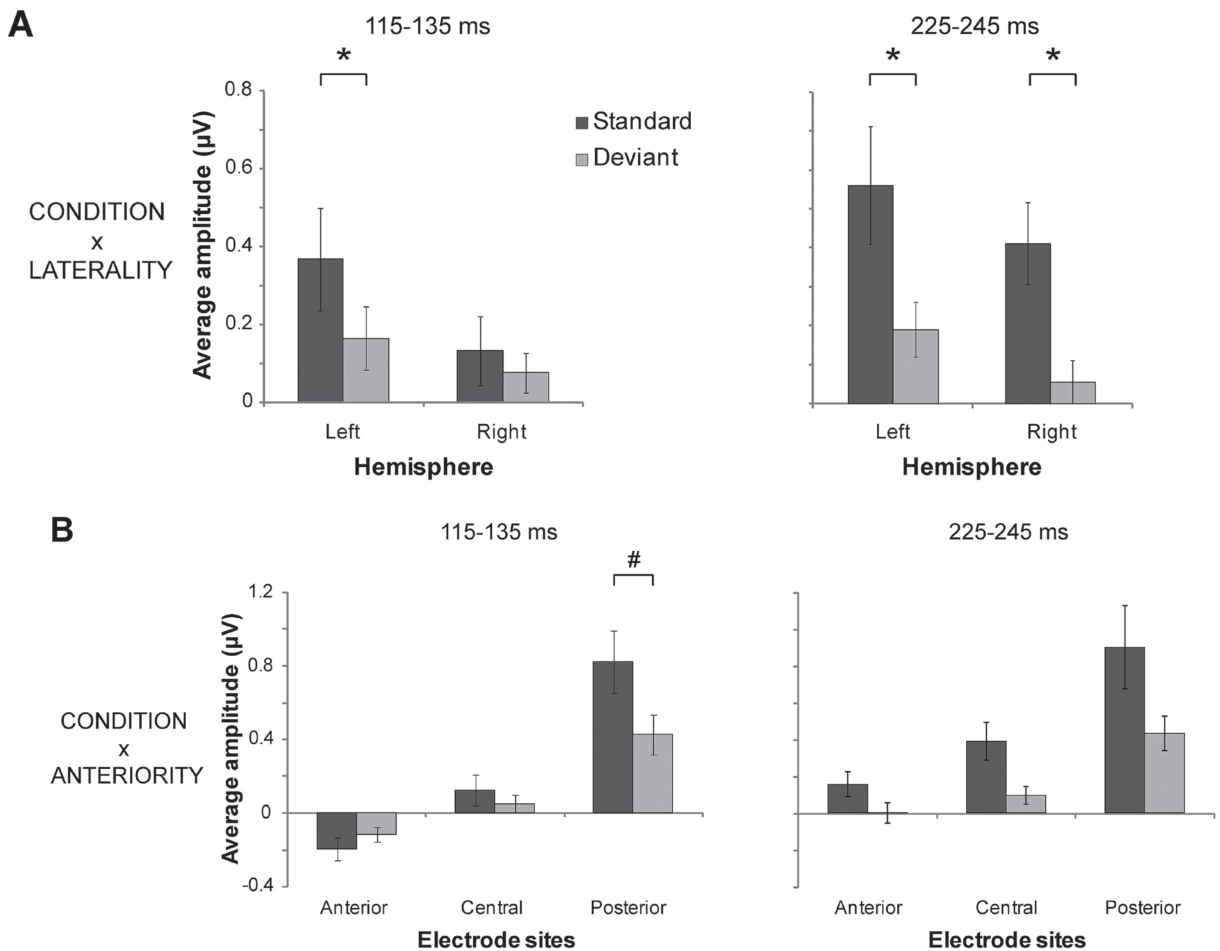

Fig. 3. Statistical interactions within the untrustworthy stimulus category. **A**: CONDITION × LATERALITY interaction revealed that standard and deviant untrustworthy faces elicited different ERPs in both analyzed time windows over the left hemisphere, and in the later time window over the right hemisphere. **B**: CONDITION × ANTERIORITY interaction revealed a tendency of more negative responses elicited by deviant compared to standard untrustworthy stimuli over posterior electrode sites. Left side: 115–135 ms time window, Right side: 225–245 ms time window. *p < .05, #p = .06

Analysis of ERP responses elicited by standard and deviant trustworthy faces in the later time window (225–245 ms) yielded a significant main effect of ANTERIORITY ($F(2,28) = 6.65$, $p < .01$, $\eta^2 = .32$), caused by more positive ERP amplitudes elicited over the posterior (occipital) electrode sites. No other effects or interactions reached significance.

Comparison of the standard trustworthy and untrustworthy ERPs in the later (225–245 ms) time window yielded a marginally significant interaction of CATEGORY × LATERALITY ($F(1,14) = 4.29$, $p = .057$, $\eta^2 = .23$). Post hoc test revealed that untrustworthy ERPs were more positive compared to trustworthy ones; and that standard untrustworthy ERPs were more positive over the left compared to the right hemisphere (Tukey HSD test: $p < .01$). Comparison of ERPs to deviant trustworthy and untrustworthy faces did not yield a statistically significant effect in the later time window.

DISCUSSION

In the present study, we demonstrated that the human brain automatically detects the regularity in a series of frequent trustworthy faces which can be violated by the presentation of rare untrustworthy faces, as indexed by the vMMN ERP component. Emotionally neutral computer-generated faces from the Trustworthiness Face Database [34] were presented to naïve subjects. Face stimuli were briefly displayed for the duration of 150 ms, which appeared to be long enough to process trustworthiness information. Recent behavioral studies also revealed that approximately 100 ms stimulus exposure is sufficient to make a final judgment about the trustworthiness of a face [35, 36].

Results of the analyses of ERPs to untrustworthy faces showed two negative intervals (115–135 ms, and 225–245 ms) of the deviant minus standard waveform, where standard and deviant responses significantly differed, which is in close agreement with results of prior vMMN studies using face stimuli [2, 21, 39]. The latency of the early (115–135 ms) time window is in line with recent studies revealing that trustworthiness of a face can be detected approximately within the first 100 to150 ms after stimulus onset, and indicates that social categorization may occur in parallel with structural encoding of the face [11, 23, 37], while latency of the late (225–245 ms) time window may also be in close agreement with previous reports, which concluded that certain aspects of automatic emotional expression processing can be identified using ERPs at approximately 200 ms after stimulus onset [21, 22]. Comparing ERP data within the identical stimulus conditions (standard or deviant) between the two categories (trustworthy vs. untrustworthy) suggested two different effects of mismatch detection and adaptation. In the earlier time window, more negative ERPs elicited by the deviant untrustworthy compared to deviant trustworthy faces indicated larger prediction error responses, whereas in the later time window, more positive ERPs elicited by the standard untrustworthy compared to standard trustworthy faces indicated less adaptation to the repetition of such faces.

The possibility of the involvement of exogenous stimulus features in the formation of the vMMN response and the usage of relatively small number of electrode poses limitations for data interpretation. We suggest that further studies should employ 1) a control task with equiprobable stimulus conditions to exclude the possibility of confounding exogenous effects and 2) high-density EEG recordings to increase the spatial resolution of scalp potentials to allow studying scalp distribution and source estimation of the vMMN response.

Results of the present study point to the fact that the repetition of trustworthy faces resulted in the formation of an automatic expectation towards trustworthiness which was violated by deviant untrustworthy faces as indexed by the evoked vMMN component. In contrary, repetition of untrustworthy faces did not build up a similar expectation towards their category, thus no vMMN was observed to the rare (deviant) trustworthy faces. The question arises then, why trustworthiness may be a perceptually more salient category compared to untrustworthiness? One possible explanation could be that humans are evolutionary adapted to detect and approach potentially beneficial (or 'good-willed') cooperators in social exchange situations [8, 20]. As the present results indicate that no automatic expectation was formed for untrustworthy faces, i.e., untrustworthiness was not processed as a stimulus category, therefore we assume that the lack of features which would define them as trustworthy may serve as the basis for their identification. This notion is in line with the idea that have been put forward by Schmidt and Cohn [28] pointing out that the lack of trustworthy facial features were considered as a sign of possible deception, and human 'cheater detection' ability was hypothesized to depend on an almost statistical knowledge of the normal (trustworthy) facial expression pattern. It was proposed earlier that facial expressions, in general, may belong to the category of so called 'honest' signals and therefore they are not typically designed to be used for deceptive purposes [38]. In addition, untrustworthy faces in our stimulus set, or in everyday social interactions as well, may exhibit higher structural variance in features that define the trustworthiness dimension, possibly as a result of various intense and highly unconcealable emotions which are involuntarily expressed on the face in case of untrustworthy intentions and subsequent (deceptive) decisions [12].

Therefore, as belonging to the default facial expression pattern, extracting trustworthiness cues from the face may be less demanding to the perceptual system, than extracting cues which are characteristic to untrustworthy faces. In concert with the above notion, recent studies also revealed that distinct features describe the morphology of trustworthy and untrustworthy faces [19, 27]. Others found that humans show clear preference for symmetric features on the face [13], and that symmetric faces elicit higher levels of adaptation in ERP responses at approx. 200 ms after stimulus onset [6]. Taken together, the present results and previous findings both converge to the notion that the human brain is biased towards the perception of trustworthiness as a natural perceptual category and readily builds pre-attentive predictions towards trustworthy faces.

In summary, the present results provide novel electrophysiological evidence that an automatic expectation can be built up for faces with trustworthy expression, which is demonstrated by an early automatic vMMN response to faces violating this expectation. As an evolutionary adaptive mechanism, fast detection and evaluation of trustworthiness-related social facial cues may serve as the basis of conscious recognition of reliable partners.

ACKNOWLEDGEMENTS

This research was realized in the frames of the TÁMOP 4.2.4. A/2-11-1-2012-0001 „National Excellence Program – Elaborating and operating an inland student and researcher personal support system convergence program" The project was subsidized by the European Union and co-financed by the European Social Fund; the SROP-4.2.2/B-10/1-2010-0029 Supporting Scientific Training of Talented Youth at the University of Pécs; the Hungarian National Science Fund (OTKA) K 101762; and the University of Pécs János Szentágothai Research Center. We thank A. Todorov and his laboratory members for their courtesy of sharing the Trustworthiness Face Database.

REFERENCES

1. Astikainen, P., Ruusuvirta, T., Wikgren, J., Korhonen, T. (2004) The human brain processes visual changes that are not cued by attended auditory stimulation. *Neurosci. Lett. 368,* 231–234.
2. Astikainen, P., Hietanen, J. K. (2009) Event-related potentials to task-irrelevant changes in facial expressions. *Behav. Brain Funct. 5,* 30.
3. Berti, S. (2011) The attentional blink demonstrates automatic deviance processing in vision. *Neuroreport 22,* 664–667.
4. Bond, C. F., Berry, D. S., Omar, A. (1994) The Kernel of Truth in Judgments of Deceptiveness. *Basic Appl. Soc. Psych. 15*, 523–534.
5. Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., Angleitner, A. (2004) Thin slices of behavior as cues of personality and intelligence. *J. Pers. Soc. Psychol. 86,* 599–614.
6. Cassia, V. M., Kuefner, D., Westerlund, A., Nelson, C. A. (2006) Modulation of face-sensitive event-related potentials by canonical and distorted human faces: The role of vertical symmetry and up-down featural arrangement. *J. Cognitive Neurosci. 18,* 1343–1358.
7. Clery, H., Andersson, F., Fonlupt, P., Gomot, M. (2013) Brain correlates of automatic visual change detection. *Neuroimage 75*, 117–122.
8. Cosmides, L., Tooby, J. (1992) Cognitive adaptations for social exchange. In: Barkow, J. E., Cosmides, L., Tooby, J. (eds) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford University Press, New York.
9. Czigler, I., Balazs, L., Winkler, I. (2002) Memory-based detection of task-irrelevant visual changes. *Psychophysiology 39,* 869–873.
10. Delorme, A., Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods 134,* 9–21.
11. Dzhelyova, M., Perrett, D. I., Jentzsch, I. (2012) Temporal dynamics of trustworthiness perception. *Brain Res. 1435,* 81–90.
12. Ekman, P. (1985) *Telling lies: Clues to deceit in the marketplace, politics, and marriage.* Norton, New York.
13. Evans, D. W., Orr, P. T., Lazar, S. M., Breton, D., Gerard, J., Ledbetter, D. H., Janosco, K., Dotts, J., Batchelder, H. (2012) Human preferences for symmetry: subjective experience, cognitive conflict and cortical brain activity. *PLoS One 7,* e38966.
14. Gayle, L. C., Gal, D. E., Kieffaber, P. D. (2012) Measuring affective reactivity in individuals with autism spectrum personality traits using the visual mismatch negativity event-related brain potential. *Front Hum. Neurosci. 6*, 334.
15. Heslenfeld, D. J. (2003) Visual mismatch negativity. In: Polich, J. (ed.) *Detection of Change: Event-Related Potential and fMRI Findings*. Kluver Academic Press, Boston, MA, pp. 41–59.

16. Kecskes-Kovacs, K., Sulykos, I., Czigler, I. (2013) Visual mismatch negativity is sensitive to symmetry as a perceptual category. *Eur. J. Neurosci. 37,* 662–667.
17. Kimura, M., Schroger, E., Czigler, I. (2011) Visual mismatch negativity and its importance in visual cognitive sciences. *Neuroreport 22,* 669–673.
18. Kimura, M., Kondo, H., Ohira, H., Schroger, E. (2012) Unintentional temporal context-based prediction of emotional faces: an electrophysiological study. *Cereb Cortex 22,* 1774–1785.
19. Kleisner, K., Priplatova, L., Frost, P., Flegr, J. (2013) Trustworthy-looking face meets brown eyes. *PLoS One 8*, e53285.
20. Kovacs-Balint, Z., Bereczkei, T., Hernadi, I. (2012) The telltale face: Possible mechanisms behind defector and cooperator recognition revealed by emotional facial expression metrics. *Brit. J. Psychol*.
21. Li, X., Lu, Y., Sun, G., Gao, L., Zhao, L. (2012) Visual mismatch negativity elicited by facial expressions: new evidence from the equiprobable paradigm. *Behav. Brain Funct. 8,* 7.
22. Liddell, B. J., Williams, L. M., Rathjen, J., Shevrin, H., Gordon, E. (2004) A temporal dissociation of subliminal versus supraliminal fear perception: an event-related potential study. *J. Cogn. Neurosci. 16,* 479–486.
23. Marzi, T., Righi, S., Ottonello, S., Cincotta, M., Viggiano, M. P. (2012) Trust at first sight: evidence from ERPs. *Soc Cogn Affect Neurosci*.
24. Naatanen, R., Paavilainen, P., Rinne, T., Alho, K. (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol. 118,* 2544–2590.
25. Oosterhof, N. N., Todorov, A. (2008) The functional basis of face evaluation. *Proc. Natl. Acad. Sci. USA 105,* 11087–11092.
26. Pazo-Alvarez, P., Amenedo, E., Lorenzo-Lopez, L., Cadaveira, F. (2004) Effects of stimulus location on automatic detection of changes in motion direction in the human brain. *Neurosci. Lett. 371,* 111–116.
27. Rezlescu, C., Duchaine, B., Olivola, C. Y., Chater, N. (2012) Unfakeable facial configurations affect strategic choices in trust games with or without information about past behavior. *PLoS One 7*, e34293.
28. Schmidt, K. L., Cohn, J. F. (2001) Human facial expressions as adaptations: Evolutionary questions in facial expression research. *Yearb Phys. Anthropol. 44,* 3–24.
29. Shen, X. B., Wu, Q., Fu, X. L. (2012) Effects of the duration of expressions on the recognition of microexpressions. *J. Zhejiang Univ. Sci. B 13,* 221–230.
30. Stagg, C., Hindley, P., Tales, A., Butler, S. (2004) Visual mismatch negativity: the detection of stimulus change. *Neuroreport 15,* 659–663.
31. Stefanics, G., Kimura, M., Czigler, I. (2011) Visual mismatch negativity reveals automatic detection of sequential regularity violation. *Front Hum. Neurosci. 5,* 46.
32. Stefanics, G., Czigler, I. (2012) Automatic prediction error responses to hands with unexpected laterality: An electrophysiological study. *Neuroimage 63,* 253–261.
33. Stefanics, G., Csukly, G., Komlosi, S., Czobor, P., Czigler, I. (2012) Processing of unattended facial emotions: A visual mismatch negativity study. *Neuroimage 59,* 3042–3049.
34. Todorov, A., Baron, S. G., Oosterhof, N. N. (2008) Evaluating face trustworthiness: a model based approach. *Soc. Cogn. Affect Neurosci. 3,* 119–127.
35. Todorov, A., Pakrashi, M., Oosterhof, N. N. (2009) Evaluating Faces on Trustworthiness after Minimal Time Exposure. *Soc. Cognition 27,* 813–833.
36. Willis, J., Todorov, A. (2006) First impressions: Making up your mind after a 100-ms exposure to a face. *Psychol. Sci. 17,* 592–598.
37. Yang, D., Qi, S., Ding, C., Song, Y. (2011) An ERP study on the time course of facial trustworthiness appraisal. *Neurosci. Lett. 496,* 147–151.
38. Zahavi, A. (1993) The fallacy of conventional signalling. *Philos. Trans R Soc. Lond. B Biol. Sci. 340,* 227–230.
39. Zhao, L., Li, J. (2006) Visual mismatch negativity elicited by facial expressions under non-attentional condition. *Neurosci. Lett. 410,* 126–131.