UNIVERSITY OF PECS

Doctoral School of Biology

Who can we trust? Cognitive neuropsychological investigations on the recognition of complex social facial expressions

PhD thesis

Zsófia Kovács-Bálint

Supervisor:

István Hernádi, PhD.

associate professor

PECS, 2013.

INTRODUCTION

Reciprocity and co-operation are adaptive social behavioural patterns, that increases the organism's fitness with the expectation that his/her counterpart will act similarly (co-operate) in a similar manner at a later time. Manifestation of these two characteristic behavioural patterns of social exchange situations may be observable among relatives and strangers, as well. Prior studies in evolutionary psychology suggest that specific cognitive mechanisms have been evolved to enable co-operators to avoid exploitation from defectors.

Defector-detector ability, theory of a cognitive brain module has been investigated from the early 80s, and investigations about the "defector facial micro-expression" began in the 90s. Research evidenced that humans remember faces of defectors better than faces of co-operators; and, that humans accurately discriminate non-co-operative faces from co-operative ones during the decision-making. On the basis of these studies we presumed that, in a social exchange situation (at the moment of cheating), defectors may feel strong, unconcealable emotions (fear, guilt, anger), so they would openly express their emotions of deceptive act. Several studies so far investigated unchangeable facial characteristics related to facial trustworthiness. However, no previous research have been addressed either the exact time resolution of the processing of 'complex social facial expressions', such as cheating and co-operating, or the brain correlates responsible for processing these complex social facial micro-expressions.

HYPOTHESIS

On the basis of earlier studies, presuming the existence of the cheater detection ability, we hypothesized that humans can read their partner's intentions from their facial expression at the moment of cheating and co-operating. According to these presumptions, our aims were to:

1. Develop and evaluate a standard face image database, which contains real defector and co-operator facial images, which we can use as stimulus in the following experiments to study complex facial expressions of deception and co-operation.
2. Reveal possible gender differences underlying complex social facial expression recognition.
3. Describe those complex defector/co-operator facial micro-expressions which appear at the moment of decision-making, by analysing facial muscle tension.
4. Investigate the time course of defector and co-operator facial expression processing, by investigating the stimulus-evoked cortical activation-patterns. We aimed to reveal that in which phase of the automatic face processing are the defector and the co-operator faces encoded.
5. Identify brain regions which may be responsible for the encoding and processing of defector and co-operator facial expressions. Seek experimental evidence on the existence and localization of the 'defector-detector' and the 'co-operator-detector' brain modules using functional magnetic resonance imaging (fMRI).
6. Evidence that humans recognize trustworthiness in an automatic manner; that trustworthy / untrustworthy facial features are processed without conscious attentional involvement (visual mismatch negativity (vMMN) component is a useful index to study automatic pre-attentive change detection processes in the brain).

Face image database development and validation

Methods

First, a custom-built computer program was developed and was used for recording the decisions of participants in a Prisoners' Dilemma Game (PDG) and to take their full-face photographs at the moment of decision making. One hundred and sixteen subjects participated in this phase; 90 of them were females (34 defectors, 56 co-operators), and 26 of them were males (11 defectors, 15 co-operators). Participants played two games: 1) decision game in a control (neutral) dilemma (whether or not to buy a new pet) and 2) decision game in a one-shot classical PDG (whether or not to co-operate with a fictional partner in a hypothetical game where 'co-operating' meant holding out on the police after a bank robbery and 'non-co-operating' meant confessing to the police). At the very moment of decision making (within a 100 ms time frame), both in the neutral and in the PDG game, a full-face web-camera photo (a neutral photo and an 'action' photo, respectively) was saved on the computer for further use as stimuli.

Results

We selected 67 from the initial 116 pairs of photos that were free from photographic deficiencies (faces that were partly covered by hair, garments, eyeglasses, or hands were excluded). The mean (±SE) age of the co-operators (n=36, 27 females) was 21.27 (±0.36) yrs., while the mean age of the defectors (n=31, 21 females) was 21.51 (±0.44) yrs. Selected photos were cropped within an oval frame and the background was masked with a black ellipse. Luminance and RGB values were adjusted to the average to standardize photos for use as images. Then, a pixel based image analysis was performed to ensure that images were the same in size, color and luminance.

Defector/co-operator recognition task

Methods

In this phase of the experiment, we tested whether the previously taken defector facial images were identifiable as defectors, and co-operator facial images as co-operators. Sixty-two healthy university students (33 females) aged 22.05 ±0.45 yrs. volunteered to participate in this experiment as lay judges. They evaluated face stimuli ('action photos') along the intensity dimension on an 11-point Likert scale. Their task was to decide whether the person shown in the image could have been a possible co-operator (1 to 5), a defector (-1 to -5) or none of those (0).

Results

We found that 17/36 (47.2%) of the co-operator images (16 females, mean score: 1.26 ±0.3), and 18/31 (58.1%) of the defector images (9 females, mean score: -0.91 ±0.14) were correctly (true-positively) identified. These true-positively identified images (n=35) were further analysed, and we found that:

1. The mean scores of the true-positively identified defector and co-operator images differed;
2. Lay judges rated male and female images equally along the defector - co-operator dimension;
3. Male judges evaluated the images with a relative bias toward the co-operator category compared to female judges, who did not show such bias;
4. Female judges gave higher scores for defector faces compared to male judges.

Defector and co-operator micro-expression encoding

Methods

To test the present hypothesis, three independent coders compared 'action' facial images (defector or co-operator) with their neutral counterparts, using the 'Facial Action Coding System' (FACS) developed by Ekman and Friesen (1976). They evaluated the entire PDG face database for facial expression metrics along 27 facial areas, termed as action units (AUs), all of which represented contraction or relaxation of one or more facial or neck muscles. Coders compared the individual AUs on each pair of 'neutral' and 'action' photos on a -5 to +5 scale using positive values for increased muscle tension, negative values for decreased muscle tension and zero for no change.

Results

Analysis of defector and co-operator micro-expressions revealed that the defectors' lips became more stretched and tightened compared to co-operators; and that co-operators opened their eyes more widely compared to defectors at the moment of decision making in the PDG (Figure 1).

Figure 1. Typical co-operator (right) and defector (left) facial expression.
For anonymity reasons, images were averaged from individual photos by morphing technique.
Bubbles indicate the approximate positions of the characteristic action units (AUs) corresponding to the facial action coding system (FACS).

Methods

Twenty-two healthy human volunteers (11 females; mean age 21.86 ±0.56 yrs.) were recruited for a single session electroencephalographic (EEG) experiment, where event related potentials (ERPs) were recorded. Stimuli consisted of 36 grayscale images of faces (12 co-operator and 12 defector facial images taken in a PDG, and 12 images with neutral facial expression), and 12 houses. Stimuli were presented in nine blocks to sustain attention and avoid repetition effect. To monitor attention, the participants' task was to detect randomly interspersed red-shaded target trials (stimuli or crosshairs).

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. Impedances were kept below 4kΩ. Recording was continuous with an analogue band-pass from 0.16 Hz to 150 Hz. EEG was digitized at 1kHz 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. EEG was analyzed in Matlab (MathWorks, Natick, MA) using the EEGLAB toolbox.

Results

Canonical face-related ERPs

1. Results revealed that co-operator, defector and neutral facial expressions evoked similar P1 amplitude values.
2. Co-operator faces evoked more negative N170 amplitude values over the right hemisphere compared to neutral faces, and more negative brain responses indicating that processing of co-operator facial expression occurs in parallel with structural encoding of the face.
3. Defector faces evoked higher P2 amplitudes compared to neutral faces over the right hemisphere (F4 and C4 electrode locations, Figure 2), which points to the fact that processing of defector faces occurs very fast and automatically, at approximately 200 ms after stimulus onset.

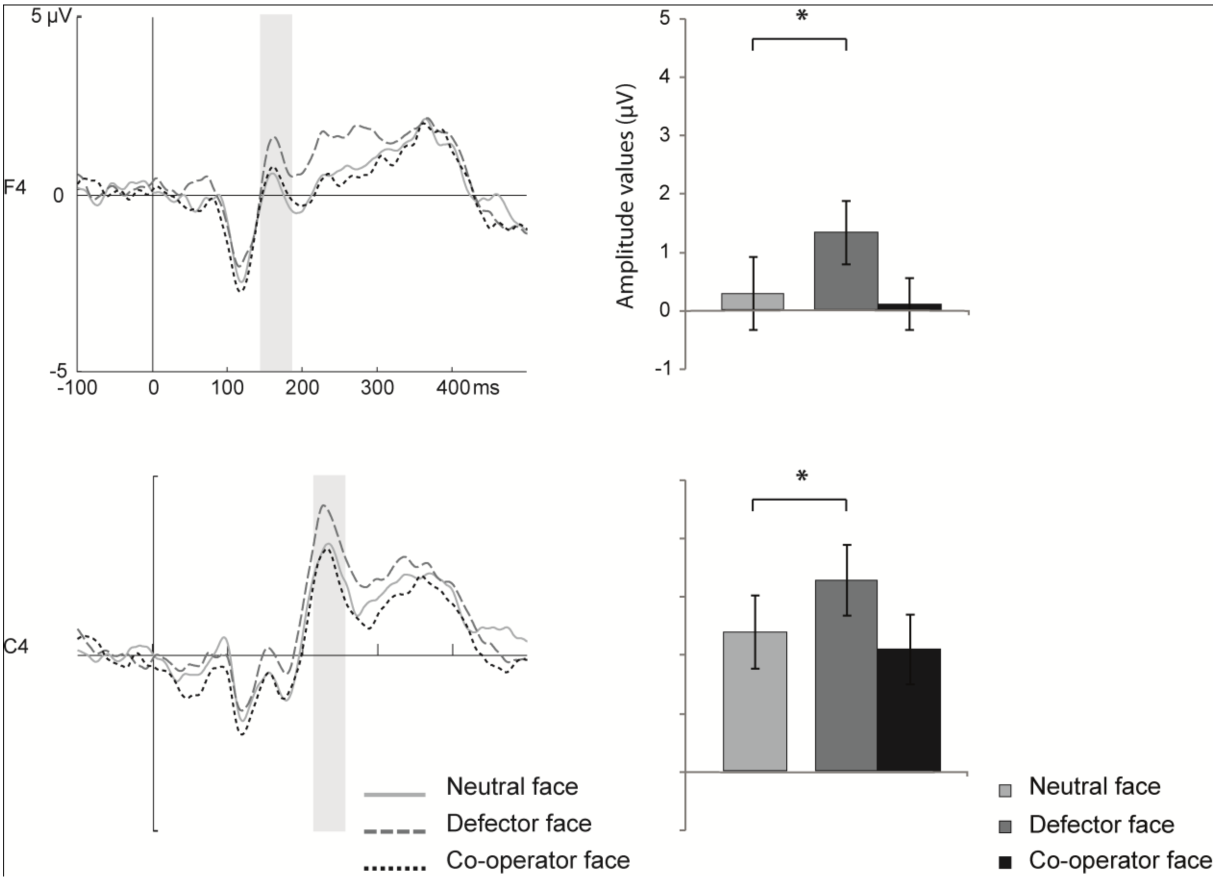

Figure 2. Defector and co-operator ERP amplitude values compared to neutral ones. Defector faces evoked higher P2 amplitude values compared to neutral ones on F4 and C4 electrode locations. *: $p<.01$

Difference wave analysis

Difference waves were calculated by subtracting ERPs elicited by neutral stimuli from ERPs elicited by defector or co-operator stimuli. Mean amplitudes around the maximum of the grand mean (±20 ms) of the P2 component were calculated. Analysing mean amplitude values of the defector-minus-neutral and the co-operator-minus-neutral difference waves, we found that defector faces elicited significantly more positive P2 amplitudes compared to co-operator faces over the C4, P4 and O2 electrode sites.

Methods

Twenty-nine right-handed university students (15 females, mean age 24 ±0.43 years) were recruited in this experiment.

Forty facial images of different identity were briefly presented to the participants in two consecutive blocks using an event-related paradigm. In the first part (control phase) of each block, 8 emotionally neutral faces and 8 houses were presented, while in the second part (recognition phase) 16 defector and 16 co-operator facial images were presented. In one block, subjects' task was to choose the defectors, and in the other block, their task was to identify co-operators. The order of the different blocks was counterbalanced across participants.

Data were collected on a 3T Siemens Magnetom TrioTim MRI scanner, using the following settings: TR/TE = 2000/36 ms, flip angle = 76°, slice thickness = 4mm, slice nr. = 23.

Data (blood oxygenation level dependent signal, BOLD signal) were analysed off-line, using Matlab (MathWorks, Natick, MA) SPM5 toolbox (Wellcome Department of Cognitive Neurology, London, UK). Contrasts were defined as: defector face > neutral face, and co-operator face > neutral face. Hypotheses were tested with the paired-sample t-tests for second level group analysis.

Results

Analysis of the defector face > neutral face contrast revealed that defector faces evoked higher BOLD activation over the left medial frontal gyrus (BA 9, Figure 3), the left cuneus (BA 17), and the left parahippocampal gyrus (BA 19).

Co-operator faces evoked higher activation over the left fusiform gyrus (BA 37), and the occipital visual areas (BA 17, 18, 19) over both hemispheres, compared to neutral faces (Figure 3).

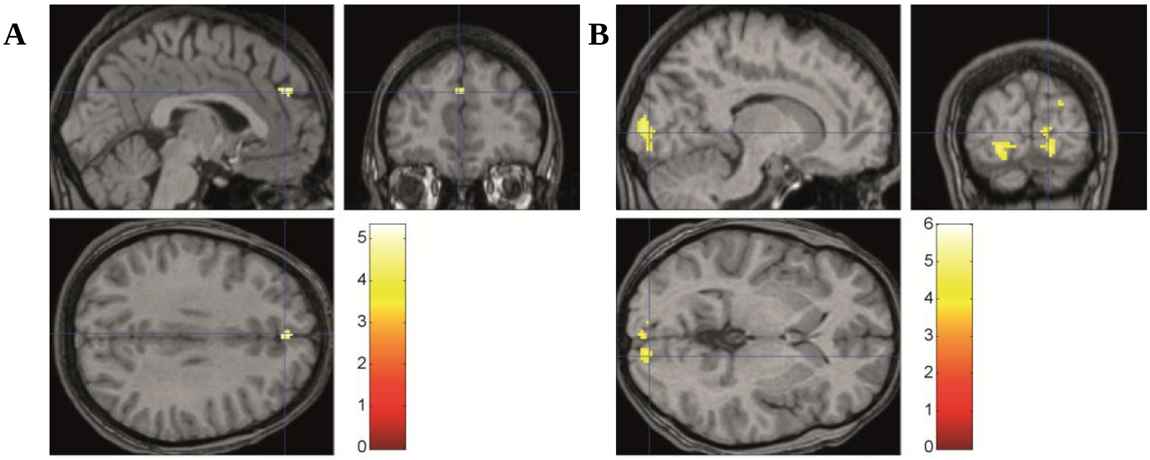

Figure 3. Brain areas activated by complex social facial expressions.

A: Defector faces evoked higher BOLD response compared to neutral faces in the left frontal lobe over area BA 9. **B:** Co-operator faces evoked higher BOLD response compared to neutral faces in both occipital lobes, over areas BA 17, 18, 19. (Level of activation is indicated by coloured dots, $p < .001$.)

CORTICAL PROCESSING OF TRUSTWORTHY/UNTRUSTWORTHY FACIAL FEATURES: A VISUAL MISMATCH NEGATIVITY STUDY

Methods

Data from 15 participants (10 females, mean age 21.27 ±0.91) were analysed in this experiment. Forty computer-generated faces of different identities (20 trustworthy and 20 untrustworthy) were used to create visual stimuli. Each stimulus consisted of four different faces (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. Two thousands stimuli were presented to the participants in two consecutive blocks, using odd-ball paradigm (1000 stimuli per block). 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 centre 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 inter-stimulus intervals (ISIs), and respond with pressing a button with their right hand, while ignoring the faces presented independently in the background.

Continuous EEG data were recorded with the same experimental arrangement as described in the previous EEG experiment 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. Mean vMMN amplitude measurements were performed around two previously identified negative peaks (between 115 and 135 ms and 225 and 245 ms, respectively) in 20 ms time windows. Mean amplitudes in both selected time windows were calculated for six electrode locations: F3, F4, C3, C4, O1 and O2, respectively.

Results

In the earlier (115-135 ms) time window, analysis of the amplitude values for ERPs evoked by standard and deviant trustworthy faces yielded no significant difference. However, untrustworthy faces evoked significant vMMN difference, more robust over the left hemisphere, and over posterior electrode locations. Post hoc tests revealed that more negative ERPs elicited by the deviant untrustworthy compared to deviant trustworthy faces indicated prediction error responses over (O1 and O2) posterior electrode sites.

Analyses in the late (225-245 ms) time window revealed that only untrustworthy faces can evoke significant vMMN response. In this time window, difference was robust over both hemispheres and it was more powerful over posterior electrode sites. More positive ERPs elicited by the standard untrustworthy compared to standard trustworthy faces indicated less adaptation to the repetition of such faces.

SUMMARY

We conducted a series of behavioural and neurophysiological (EEG and fMRI) experiments to investigate the psychophysiological mechanisms that underlying the perception of complex social facial expressions. It was hypothesized that that 1) deception may be associated with a momentary facial expression which is evoked by strong negative emotions appearing on the face and 2) co-operation - as being a natural decision category - lacks such facial expressions.

First, we **developed a standardized face database**, which contained defector and co-operator facial images and was appropriate for further use in the EEG and fMRI experiments.

Next, facial expression recognition was tested and in agreement with the previously established 'cheater detection theory', we hypothesized that humans are able to recognize defector partners in social encounters based on momentary changes of their facial expressions. Results indicated that **female observers recognize defector faces more confidently compared to males**; and we also evidenced that **males evaluated the facial images with a relative bias toward the co-operator category** compared to females. Thanks to the well-established facial expression analysis method (FACS) we characterized defector facial micro-expressions. Results revealed that **defectors but not co-operators closed their upper eyelids as if they were blinking, and also depressed and tightened their (lower) lips**.

To study the time course of cortical brain processes and the regions activated by defector- and co-operator facial expressions, we conducted event-related EEG and fMRI experiments, respectively. The EEG experiment revealed that the **processing of co-operator facial expression occurs in parallel with structural encoding of the face, with higher activation over the right hemisphere (indicated by the N170 ERP component)**; and that **processing of defector faces occurs automatically, at approximately 200 ms after stimulus onset (indicated by the P2 ERP component)**, immediately after the structural encoding of the face. Different processing of defector and neutral facial expressions over frontal electrode sites evidenced that the encoding of **defector faces may demand higher attentional involvement**.

Results of the fMRI experiment revealed that the **processing of co-operator facial expressions resulted in higher activation in right the occipital face area (OFA), the right fusiform face area (FFA),** while the **processing of defector facial expressions activated the left pre-frontal cortex (PFC)**. Current results led to the conclusion that in case of co-operator faces, greater activation in the OFA and the FFA regions - which are parts of the core face

processing system – means that co-operator faces were processed as a natural perceptual category, which is generally preferred compared to defector faces. In contrary, the PFC, which plays crucial role in social perception and in regulation of action planning and execution, actively participates in the allocation of greater amounts of attention to defector faces.

Results of the vMMN experiment support the existence of the previously supposed 'co-operator-detector module' in the brain. Here we evidenced **the formation of an unconscious expectation towards the trustworthiness dimension** which was violated by deviant untrustworthy faces but not vice versa, as indexed by the vMMN component. In line with previous data, the present results also point to the fact that **the lack of trustworthy facial features on the face may be a sign of possible deceptive action**, and that successful defector detection may depend on the violation of the prior "statistical" knowledge of the normal (co-operator) facial expression pattern.

In summary, the present series of experiments, for the first time, evidenced the difference between micro-expressions of defector- and co-operator facial expressions; and that the processing of defector and co-operator facial expressions is a very fast and automatic process, apparently supervised by the OFA, the FFA and the PFC cortical regions.

Recent studies on brain processing revealed that key aspects of emotional facial expressions may induce an automatic impression formation which takes place approximately within the first 100 ms after stimulus onset (presumably due to the involvement of limbic cortical regions, such as the amygdala the cingulate cortices); while the results of the present experiments revealed that the processing of complex social facial expressions such as cheating or cooperation occurs later, after the initial processing of structural facial characteristics and facial expressions; and may involve higher cortical activity with frontal lobe control. These notions led to the conclusion that the previously well-established cognitive model of face perception (Bruce and Young, 1986), which also suggests the time course of general face processing, should be further modified to a so-called 'socio-cognitive face perception model' including key aspects of the processing of the socially relevant facial features and complex social facial expressions.

PUBLICATIONS

Publications related to the thesis

Peer-reviewed articles:

1. **Kovács-Bálint, Zs.**, Bereczkei, T., Hernádi, I. 2012: The telltale face: Possible mechanisms behind defector and cooperator recognition revealed by emotional facial expression metrics. Brit J Psychol, e-pub: 21.11.2012, DOI: 10.1111/bjop.12007. **IF: 2.103**
2. **Kovács-Bálint, Zs.**, Stefanics, G., Trunk, A., Hernádi, I. 2014: Automatic detection of trustworthiness of the face: a visual mismatch negativity study. Acta Biol Hun, 65(1), In press. **IF: 0.504**
3. **Kovács-Bálint, Zs.,** Hernádi I. 2013: The time course of face perception in the human brain. From structural coding to social cognition. Acta Neurob Exp, Under revision.

Oral and poster presentations:

1. **Kovács-Bálint, Zs.**, Stefanics, G., Hernádi, I. 2009: A csalódetekció fiziológiai hátterének humán vizsgálata. Szóbeli előadás: Biológus Doktoranduszok Konferenciája, Pécs.
2. **Kovács-Bálint, Zs.**, Pachner, O., Stefanics, G., Hernádi, I. 2009: Psychophysiological background of cheater detection. Poster presentation: 73rd Meeting of the Hungarian Physiological Society. Acta Physiol Hun, 97(1), 118.
3. **Kovács-Bálint, Zs.**, Stefanics, G., Trunk, A., Hernádi, I. 2010: Subconscious recognition of cheater faces in a visual discrimination task. Poster presentation: IBRO International Workshop. Front Neurosci, Doi: 10.3389/conf.fnins.2010.10.00163
4. **Kovács-Bálint, Zs.**, Trunk, A., Stefanics, G., Hernádi, I. 2010: Behavioral correlates of cheater detection in humans. Poster presentation: 7th FENS, Amsterdam.
5. **Kovács-Bálint, Zs.**, Trunk, A., Stefanics, G., Hernádi, I. 2011: Evidence for automatic cheater detection as indexed by event related brain potentials. Poster presentation: 13th Conference of the Hungarian Neuroscience Society. Front Neurosci, Doi: 10.3389/conf.fnins.2011.84.00166
6. **Kovács-Bálint, Zs.**, Trunk, A., Stefanics, G., Hernádi, I. 2011: Who can we trust? Early visual processing of differential facial expressions related to trustworthiness: An ERP study. Poster presentation: 75th Meeting of the Hungarian Physiological Society. Acta Physiol Scand, 202(S684), 47.

7. **Kovács-Bálint, Zs.**, Trunk, A., Stefanics, G., Hernádi, I. 2012a: Early visual processing of socio-economically relevant facial expressions related to deception: An event-related brain potential study. Poster presentation: 8th FENS, Barcelona.
8. **Kovács-Bálint, Zs.**, Trunk, A., Stefanics, G., Hernádi, I. 2012b: Processing of socio-economically relevant facial expressions as revealed by event-related brain potentials. Poster presentation: IBRO International Workshop. Clin Neurosci, 65(1.suppl), 38.
9. **Kovács-Bálint, Zs.**, Deák, A., Papp, P., Perlaki G., Orsi, G., Hernádi, I., Bereczkei, T. 2013: Representation of cooperative intention in the human brain based on facial expression. Poster presentation: 14th Conference of the Hungarian Neuroscience Society. ISBN: 978-963-88224-2-0
10. **Kovács-Bálint, Zs.**, Deák, A., Papp, P., Perlaki G., Orsi, G., Hernádi, I., Bereczkei, T. 2013: Processing facial expressions of cooperation and deception activates different cortical regions: an fMRI study. Poster presentation: 18th Meeting of the European Society of Cognitive Psychology, Budapest.

Other publications

Peer-reviewed articles:

1. **Kovács-Bálint, Zs.,** Csathó, Á., László, JF., Juhász, P., Hernádi, I. 2011: Exposure to an inhomogeneous static magnetic field increases thermal pain threshold in healthy human volunteers. Bioelectromagnetics, 32(2), 131-139. **IF: 1,842**
2. Trunk, A., Stefanics, G., Zentai, N., **Kovács-Bálint, Zs.**, Thuróczy, Gy., Hernádi, I. 2013: No effects of a single 3G UMTS mobile phone exposure on spontaneous EEG activity, ERP correlates, and automatic deviance detection. Bioelectromagnetics, 34(1), 31-42. **IF: 2,021**

Oral and poster presentations:

1. **Kovács-Bálint, Zs.,** Csathó, Á., Pachner, O., László, J., Hernádi, I. 2009: Effects of inhomogeneous static magnetic field on human thermal pain threshold. Poster prsentation: 12th Meeting of the Hungarian Neuroscience Society. Front Sys Neurosci, Doi: 10.3389/conf.neuro.01.2009.04.233
2. Stefanics, G., Trunk, A., **Kovács-Bálint, Zs.**, Thuróczy, Gy., Hernádi, I. 2010: The effect of thirty-minute 3G mobile phone exposure on auditory evoked potentials and the MMN response. Poster presentation. Int J Psychophysiol, 77(3), 327-328.

3. Trunk, A., Stefanics, G., **Kovács-Bálint, Zs.**, Thuróczy, Gy., Hernádi, I. 2010: The effect of 3G EMF exposure on auditory evoked potentials and automatic deviance detection: an EPR study. Poster presentation: IBRO International Workshop. Front Neurosci, Doi: 10.3389/conf.fnins.2010.10.00171
4. Trunk, A., Stefanics, G., **Kovács-Bálint, Zs.**, Thuróczy, Gy., Hernádi, I. 2010: The effect of a single thirty-minute 3G mobile phone irradiation on auditory evoked potentials. Poster presentation: 7th FENS, Amsterdam.
5. Trunk, A., Stefanics, G., **Kovács-Bálint, Zs**., Zentai, N., Thuróczy, Gy., Hernádi, I. 2011: The effect of single 30 minute long 3G EMF exposure on auditory evoked potentials, automatic deviance detection and spontaneous EEG. Poster presentation: 13th Conference of the Hungarian Neuroscience Society. Front Neurosci, Doi: 10.3389/conf.fnins.2011.84.00061
6. **Kovács-Bálint, Zs.**, Trunk, A., Stefanics, G., Hernádi, I. 2012: Kísérlettervezés a neuropszichológiában. Oral presentation: “Test és Lélek” – Határterületek a Neuropszichológiában és Pszichoterápiában Konferencia, Pécs.

Lecture note (e-book):

1. Hernádi, I., Dénes, V., **Kovács-Bálint, Zs.** 2009: Fiziológiai gyakorlatok és mérések a pszichológiában. Egyetemi elektronikus jegyzet.

Scientometrics

Peer-reviewed articles: 4;

Cumulative impact factor: 6.5;

Citation index: 7 (independent: 4)