

doi: 10.1093/scan/nsz007

Advance Access Publication Date: 25 January 2019 Original article

Theta resting EEG in the right TPJ is associated with individual differences in implicit intergroup bias

Bastian Schiller,^{1,3,4,*} Lorena R.R. Gianotti,^{2,3,*} Thomas Baumgartner,^{2,3} and Daria Knoch^{2,3}

¹Department of Psychology, Laboratory for Biological and Personality Psychology, University of Freiburg, D-79104, Freiburg, Germany, ²Institute of Psychology, Department of Social Psychology and Social Neuroscience, University of Bern, CH-3012, Bern, Switzerland, ³Department of Psychology, Social and Affective Neuroscience, University of Basel, CH-4055, Basel, Switzerland, and ⁴Freiburg Brain Imaging Center, University Medical Center, University of Freiburg, D-79104, Freiburg, Germany

Correspondence should be addressed to Bastian Schiller, Department of Psychology, Laboratory for Biological and Personality Psychology, University of Freiburg, Stefan-Meier-Strasse 8, D-79104 Freiburg i. Br, Germany. E-mail: schiller@psychologie.uni-freiburg.de.
*Bastian Schiller and Lorena R.R. Gianotti contributed equally to this work.

Abstract

Why are some people more biased than others in their implicit evaluations during social interaction? The dispositional determinants of individual differences in implicit intergroup bias are poorly understood. Here, we explored whether such variability might be explained by stable neural traits. For that purpose, we used the source-localized resting electroencephalograms of 83 members of naturalistic social groups to explain their bias in an in-/outgroup implicit association test. Lower levels of resting theta current density in the right temporo-parietal junction (TPJ) were associated with stronger implicit intergroup bias and explained unique variability in bias beyond relevant personality questionnaires. These findings demonstrate the added value of the neural trait approach in predicting inter-individual differences in implicit social cognition. Given that low levels of resting theta current density during wakefulness likely reflect increased cortical activation, our results suggest that individuals with an efficiently working right TPJ possess capacities to mediate specific cognitive processes that predispose them towards stronger implicit intergroup bias. As the human species has evolved living in distinct social groups, the capacity to quickly differentiate friend from foe became highly adaptive and might thus constitute an essential part of human nature.

Key words: implicit intergroup bias; resting EEG; right TPJ; individual differences; neural traits

Introduction

People differ in the degree to which their social interactions are biased by their counterpart's group affiliation, for example her or his political party, sport club, religion or nation (for reviews see Hewstone *et al.*, 2002; Cikara and Van Bavel, 2014). Such individual differences in interactions with in- and outgroup members are often driven by differences in early, automatically

favorable or unfavorable evaluations of others (Greenwald et al., 2009; Kubota et al., 2013), termed 'implicit attitudes' (Eagly, 1998). To obtain more complete understanding of why an individual displays more or less intergroup bias, it is therefore crucial to illuminate why an individual possesses more or less biased implicit intergroup attitudes, i.e. implicit intergroup bias. Of note, implicit intergroup bias is already observed in children around age 6 and seems to remain strikingly invariant during

Received: 23 March 2018; Revised: 15 January 2019; Accepted: 21 January 2019

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

development (Dunham et al., 2008; Dunham et al., 2013), indicating its trait-like character. Surprisingly, given the plethora of tests to measure variability in implicit intergroup bias (for reviews see De Houwer et al., 2009; Nosek et al., 2011), few studies have addressed its dispositional determinants (but see Pratto and Shih, 2000; Cunningham et al., 2004; Rowatt et al., 2005; Bergh et al., 2012). An ideal way to shed light on this issue is provided by neuroscientific methods enabling an individual's neural traits to be objectively quantified, which in turn can help illuminate the sources of variability in the construct of interest (Hahn et al., 2015a; Nash et al., 2015). Therefore, in the present study, we sought to explore whether neural traits—dispositional brain-based characteristics—might drive individual differences in implicit intergroup bias.

Measuring an individual's electroencephalographic activity at rest represents an ideal neural trait measure because it is stable (Dunki et al., 2000; Napflin et al., 2007; Cannon et al., 2012), heritable (Smit et al., 2005; Zietsch et al., 2007; de Geus, 2010) and unique to the individual (Dunki et al., 2000; Napflin et al., 2007). Moreover, due to its high temporal resolution, resting electroencephalography (EEG) reveals the brain's oscillatory activity at rest (for reviews see Buzsaki and Draguhn, 2004; Lopes da Silva, 2013). To our knowledge, while resting EEG has successfully been used to explain individual variability in diverse phenotypes in healthy participants (Harmon-Jones et al., 2010; Knoch et al., 2010; Gianotti et al., 2012; Baumgartner et al., 2013a; Schiller et al., 2014b; Hahn et al., 2015b; Gianotti et al., 2018a; Gianotti et al., 2018b), so far, no study has employed it in predicting implicit intergroup bias. Thus, we capitalized here on resting EEG in order to effectively capture dispositional differences in neural baseline activation that can be related to individual differences in implicit intergroup bias.

To our knowledge, so far no study has investigated which regions' baseline activation relates to implicit intergroup bias. Functional neuroimaging studies have identified brain activation during intergroup bias in key regions such as the anterior cingulate cortex (ACC), bilateral insula, medial prefrontal cortex (MPFC), bilateral orbitofrontal cortex (OFC) and bilateral temporo-parietal junction (TPJ; Amodio, 2014; Baumgartner et al., 2012; Cikara and Van Bavel, 2014; Kubota et al., 2012; Molenberghs, 2013; Strombach et al., 2015). However, functional imaging studies, although indispensable, do not permit causal inferences about the role of brain regions in intergroup bias, because the observed neural activations could simply be an epiphenomenon or a consequence, and not necessarily the cause of the biased perception, judgment or behavior. In contrast, brain modulation studies interfere non-invasively with the activity of specific areas in the human cortex (e.g. Candidi et al., 2015; Korb et al., 2015) and allow researchers to draw causal conclusions about the impact of the stimulated brain region on intergroup bias (Marini et al., 2018). For example, there is evidence that temporarily disrupting the right TPJ by means of transcranial magnetic stimulation caused a decrease in behavioral intergroup bias (Baumgartner et al., 2014). Other brain modulation studies have directly investigated the effects of experimentally modulating brain activity on implicit intergroup bias by targeting the bilateral anterior temporal lobes (aTLs) and (lateral and medial) prefrontal areas. While disrupting activation of the aTLs diminished implicit intergroup bias (Gallate et al., 2011; Wong et al., 2012), disrupting activation of prefrontal areas raised implicit intergroup bias (Cattaneo et al., 2011).

In the present study, we registered the task-independent resting EEGs of 84 participants before they took an implicit association test (IAT) that determined their individual degree of implicit intergroup bias (Greenwald et al., 1998; see Material and methods). Given that disrupting activity in the right TPJ and the bilateral aTLs diminishes intergroup bias, we hypothesized that baseline activation of these regions would relate positively to implicit intergroup bias. Given that disrupting activity in prefrontal areas raises bias, we hypothesized that baseline activation of prefrontal areas would relate negatively to implicit intergroup bias. As those studies used disruptive brain stimulation protocols known to primarily affect EEG slow-wave oscillations (Wozniak-Kwasniewska et al., 2013), we expected to find correlations between baseline activation and bias in the slow-wave EEG frequency bands. In addition, we explored whether the baseline activation of brain regions in which correlational activation had been observed during intergroup bias (insula, ACC, OFC) would also relate to implicit intergroup bias.

Materials and methods

Participants

We recruited 84 right-handed and German-speaking participants from the University of Basel, Switzerland. No participant had a current or previous history of neurological or psychiatric disorders and alcohol or drug abuse. One participant had to be excluded from further analysis because of excessive EEG artifacts, leaving a sample of 83 participants (37 soccer fans: 19 females; 46 political supporters: 19 females) with a mean age of 21.9 years (s.d. = 3.0 years, range: 21-48 years). To explore the sources of individual differences in implicit intergroup bias, we used naturally occurring social groups. In an online questionnaire, we asked a large sample of students about their interests in several domains (e.g. arts, music, politics, religion, soccer) in order to keep participants blind to the purpose of the study before the experiment. Among pre-screened participants, we recruited those participants who had, on a scale from 1 (very weak) to 5 (very strong), at least medium (=3) self-reported interest in soccer (N = 37) or in politics (N = 46), because previous studies using these groups have reported strong intergroup biases (Hein et al., 2010; Schiller et al., 2014a). We validated each participant's strength of identification with his or her favored soccer club or political party by using a modified version of the Sport Spectator Identification Scale (5-point Likert scale; Wann and Branscombe, 1993) that was collected online after the laboratory experiment (see Procedure). On average, participants showed a medium to strong identification with their group (M = 3.39, s.d. = 0.58). There were no significant differences between soccer fans and political supporters (soccer fans: M = 3.50, s.d. = 0.59; political supporters: M = 3.35, s.d. = 0.55; soccer fans vs political supporters: $T_{(80)} = 1.15$, P > 0.20). Participants received 30 Swiss francs (CHF 1 = \$US) for participation.

Procedure

The study was approved by our local ethics committee and conducted according to the Declaration of Helsinki. All procedures were carried out with the adequate understanding and informed consent of the participants. After placement of the EEG electrodes, participants were seated comfortably in a dimly lit, quiet room, with intercom connection to the experimenters. They were instructed that the EEG was to be recorded during resting with open or closed eyes. The resting EEG protocol consisted of 20-s eyes open followed by 40-s eyes closed, repeated five times. Such a protocol guarantees minimal fluctuations in participants' vigilance state. The instructions about eye opening/closing were given via intercom. Data analysis was based on the 200-s eyes-

After the resting EEG recording, participants took the IAT (for an analysis of evoked-response data recorded while taking the IAT, see Schiller et al., 2016). The resting EEG recording was always made before participants took the IAT in order to minimize variance in cognitive state by standardizing preexperimental procedure for all participants (van Diessen et al., 2015). Finally, several weeks after the resting EEG and IAT measurement, we measured personality traits (Social Dominance Orientation, Jost and Thompson, 2000; Sport Spectator Identification Scale, Wann and Branscombe, 1993; Moral Foundations Questionnaire, Graham et al., 2011) that have been related to individual differences in intergroup bias (Pratto and Shih, 2000; Hein et al., 2010; Smith et al., 2014).

IAT

Using the IAT, we measured a participant's bias in implicit intergroup attitudes by determining how strongly participants automatically associate their in- and outgroup with positive and negative valence. Participants were required to correctly and quickly classify words from four categories: ingroup (e.g. names of soccer players on the favored soccer team), outgroup (e.g. names of soccer players on the rival soccer team), positive (e.g. 'love') and negative (e.g. 'death'). For each participant, we adapted the in- and outgroup words based on his or her preferred social group (see Schiller et al., 2016 for a complete list of stimuli used for each social group). The words appeared in the middle of a PC screen in black letters against a white background, and participants were to assign the words as fast as possible by pressing one of two response keys with the left and right index finger, respectively. The rules of category-response assignments changed from block to block, and the categories were presented throughout the block in the upper left and right hand corner of the screen. The IAT contained 10 blocks (364 trials in total). In the first two blocks (each 10 trials), participants learned to classify positive vs negative words and ingroup vs outgroup words, respectively. In the third 'congruent' block (76 trials), participants had to press one key when ingroup or positive words appeared, while they had to press another key when outgroup or negative words were shown. In the fourth block response, assignments for positive and negative words were reversed (10 trials), so that in the fifth 'incongruent' block (76 trials), ingroup and negative words shared the same response key, while outgroup and positive words shared another response key. After these first five blocks, participants had to do another five-block IAT, where the order of the congruent and incongruent block was switched. Because we were interested in inter-individual differences, we used a fixed order of stimulus presentation for all participants (Friese et al., 2007; Raccuia, 2016), thereby keeping task-switching costs constant (Mierke and Klauer, 2001). In each trial, the word was presented for 1500 ms, followed by a screen where only the category labels were shown with a randomly jittered duration ranging from 2000 to 2200 ms, resulting in a mean stimulus onset asynchrony of 3600 ms.

Analysis of behavioral data

For each subject, we calculated mean reaction time (RT) and accuracy for the incongruent and congruent trials, respectively. To obtain a measure of the strength of implicit intergroup bias for each subject, we calculated the D score using the improved scoring algorithm (Greenwald et al., 2003). This score is calculated by dividing the RT difference between incongruent and congruent trials by the pooled s.d. in these trials, thus adjusting for each subject's latency variability. Positive D scores indicate a strong association between the ingroup and positive valence and/or a strong association between the outgroup and negative valence.

Resting EEG recording and pre-processing

We recorded the EEG during rest with a Biosemi ActiveTwo system from 64 Ag-AgCl active electrodes according to the 10-10 system montage (Nuwer et al., 1998). The EEG was on-line band-pass filtered between 0.1 and 100 Hz, and the data were digitalized at a sampling rate of 500 Hz. The signals were referenced on-line to the common mode sense, while driven right leg served as ground. Horizontal and vertical electro-oculographic signals were recorded with two additional electrodes at the left and right outer canthi and one electrode at the left infraorbital. Pre-processing was done using Brain Vision Analyzer (Version 2.0.1.391; Brain Products GmbH). Eye-movement artifacts were corrected using independent component analysis. EEG data containing muscle, movement and/or technical artifacts were marked; noisy channels were linearly interpolated. All artifactfree data were parsed into 2-s epochs for analysis. On average, 154 s (s.d. = 42) of EEG data were available per participant from the 200 s of eyes-closed resting condition.

A Fast Fourier Transformation (using a square window) was applied to each epoch and channel to compute the power spectra with 0.5-Hz resolution. Each channel's spectra were averaged over all epochs for each participant. Absolute power spectra were integrated for the following seven independent frequency bands (Kubicki et al., 1979): delta (1.5-6 Hz), theta (6.5-8 Hz), alpha1 (8.5-10 Hz), alpha2 (10.5-12 Hz), beta1 (12.5-18 Hz), beta2 (18.5-21 Hz) and beta3 (21.5-30 Hz). Finally, standardized low-resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, 2002) was applied to estimate the intracerebral electrical sources that generated the scalp-recorded activity, separately for each EEG frequency band. The sLORETA method is a properly standardized discrete, 3D-distributed, linear, minimum norm inverse solution. The particular form of standardization used in sLORETA endows the tomography with the property of exact localization to test point sources, yielding images of standardized current density with exact localization, albeit with low spatial resolution (i.e. neighboring neural sources will be highly correlated). sLORETA has recently been validated in several simultaneous EEG/functional magnetic resonance imaging (fMRI) studies (Mobascher et al., 2009a; Mobascher, et al., 2009b) and in an EEG localization study for epilepsy (Rullmann et al., 2009). In the current implementation of sLORETA, computations were made in a realistic head model using the MNI152 template (Mazziotta et al., 2001), with the 3D solution space restricted to cortical gray matter, as determined by the probabilistic Talairach atlas (Lancaster et al., 2000). The intracerebral volume is partitioned in 6239 voxels at 5-mm spatial resolution. Thus, sLORETA images represent the standardized electric activity at each voxel in neuroanatomic Montreal Neurological Institute (MNI) space as the exact magnitude of the estimated current density. Using the automatic regularization method in the sLORETA software, we chose the transformation matrix with the signal-to-noise ratio set to 10. To reduce confounds that have no regional specificity, for each participant, sLORETA images were normalized to a total power of one and then log-transformed before statistical analyses.

Statistical analyses

As a 'first step', a voxel-wise correlation approach was taken to identify brain regions whose baseline activations correlate with the IAT D score, separately for each EEG frequency band. For our voxel-by-voxel Pearson correlation analyses, we created a priori regions of interest (ROIs). In brain regions whose causal role in behavioral and implicit intergroup bias is known, we created 15-mm spheres centered on MNI coordinates from target regions in brain stimulation studies [Baumgartner et al., 2014: right TPJ: x = 57, y = -60, z = 30; left TPJ, x = -45, y = -60, z = 21; Gallate et al., 2011; Wong et al., 2012: left aTL: x = -52, y = 6, z = -26; right aTL: x = 54, y = 8, z = -26; Cattaneo et al., 2011: left MPFC: x = -15, y = 52, z = 13; right MPFC: x = 15, y = 52, z = 13; left dorsolateral prefrontal cortex (DLPFC): x = -47, y = 16, z = 48; right DLPFC: x = 47, y = 16, z = 48]. Additionally, using all voxels labelled accordingly in sLORETA (Pascual-Marqui, 2002), we created anatomical ROIs in the bilateral insula, ACC and bilateral OFC (Kubota et al., 2012; Molenberghs, 2013; Amodio, 2014; Cikara and Van Bavel, 2014). For analyzing the relationship between implicit bias and neural baseline activation, we created one large ROI from the abovementioned ROIs and corrected for multiple testing in all of these 741 voxels via a non-parametric randomization approach (Nichols and Holmes, 2002). That approach was taken to estimate empirical probability distributions and the corresponding critical probability thresholds corrected for multiple comparisons. As a 'second step', we checked for moderating effects of 'gender' or 'group type' (soccer vs politics) on the relationship between neural baseline activation and implicit intergroup bias. For that purpose, we calculated moderation analyses using the PROCESS procedure for SPSS (Hayes, 2013). In these analyses, we predicted implicit bias by either 'gender' or 'group type', by any ROI that significantly predicted bias and by the interaction effect of both factors ('gender * ROI' or 'group type * ROI'). A significant interaction effect would indicate that 'gender' or 'group type' moderates the relationship between neural baseline activation and bias. As a 'third step', we checked whether neural baseline activation might explain unique variability in implicit intergroup bias beyond relevant personality traits. For that purpose, we added any ROI that significantly predicted bias as predictor in the second step of a regression analysis, in which relevant personality traits (Social Dominance Orientation, Jost and Thompson, 2000; Sport Spectator Identification Scale, Wann and Branscombe, 1993; Moral Foundations Questionnaire, Graham et al., 2011) have been added as predictors in the first step. We then tested whether the added predictor would significantly increase the explained variability in implicit intergroup bias.

Results

Behavioral results

First, we checked whether, overall, there was a significant implicit intergroup bias. Participants performed at 95% accuracy in the incongruent condition and at 98% accuracy in the congruent condition [errors incongruent: M = 14.06, s.d. = 10.00; errors congruent: M = 6.84, s.d. = 6.40; T(82) = 9.16, P < 0.001, $ETA^2 = 0.51$]. RTs in incongruent trials (M = 853.10 ms, s.d. = 93.56 ms) were significantly longer than in congruent trials [M = 734.87 ms, s.d. = 95.83 ms; T(82) = 14.88, P < 0.001,

95% BCa CI (102.42, 134.02), ETA² = 0.73]. This RT difference resulted in a significantly positive D-score, indicating that on average, participants possessed biased implicit intergroup attitudes [M = 0.61; s.d. = 0.36, T(82) = 15.72, P < 0.001, 95% BCa CI (0.54, 0.69), ETA² = 0.75]. We did, however, observe considerable individual variability in this implicit intergroup bias. According to psychological conventions for effect sizes (Cohen, 1977; see Nosek and Sriram, 2007, regarding the relationship between Cohen's d and the IAT D score), 11 participants displayed little or no bias (D <= 0.20), 17 participants displayed weak bias (0.20 < D <= 0.49), 23 participants displayed medium bias (0.49 < D <= 0.74) and 32 participants displayed strong bias (D > 0.74).

Brain results

In the 'main analysis', we checked whether neural baseline activation in the ROIs would relate to individual differences in implicit intergroup bias. Using sLORETA to estimate intracerebral sources underlying scalp-recorded resting EEG, we found that in the theta frequency band (6.5-8 Hz), there were 23 voxels showing significant negative correlations between current density and the D-score (P < 0.05, corrected for multiple testing; see Figure 1). All of these voxels fell into one cluster in the BAs 22, 39 and 40 in the right TPJ ROI (MNI coordinates of peak voxel: x = 60, y = -60, z = 20, superior temporal gyrus, BA 22). The significant negative correlation between current density within the right TPJ in the theta frequency band (i.e. averaged current density across significant voxels of the right TPJ) and implicit intergroup bias was r = -0.34, P = 0.002, 95% BCa CI [-0.14, -0.52](when current density was averaged across all voxels of the right TPJ ROI, the correlation was r = -0.30, P = 0.006, 95% BCa CI [-0.10, -0.48]; following a reviewer's suggestion, we additionally confirmed our results using 10-mm spheres centered on MNI coordinates from target regions in brain stimulation studies). The statistical robustness of this main finding was corroborated by the fact that three voxels survived correction for multiple testing in the whole brain (P < 0.05). Meng's tests (Meng et al., 1992) for dependent correlations confirmed the laterality effect: The correlation between the implicit intergroup bias and current density in the right TPJ was significantly stronger than the correlation between the implicit intergroup bias and current density in the homologous area in the left TPJ (Z = -3.11, P = 0.001). No other voxel in either frequency band revealed significant correlations when correcting for multiple testing.

In an 'additional analysis', we checked for moderating effects of gender and group type on the relationship between right TPJ's baseline activation and implicit intergroup bias. Including both variables as moderators in the regression analysis revealed that neither the interaction effect of 'gender x theta activity in right TPJ' nor the interaction effect of 'group type x theta current density in right TPJ' was significant (both P > 0.20). Thus, the negative correlation between current density within the right TPJ in the theta frequency band and implicit intergroup bias applies across members of two different social groups (i.e. soccer fans, political supporters) as well as across the genders.

Finally, we explored whether current density within the right TPJ in the theta frequency band would explain unique variability in implicit intergroup bias beyond relevant personality traits (Social Dominance Orientation, Jost and Thompson, 2000; Sport Spectator Identification Scale, Wann and Branscombe, 1993; Moral Foundations Questionnaire, Graham et al., 2011). Indeed, when the right TPJ's theta current density was added as a predictor in the second step of a regression analysis, it

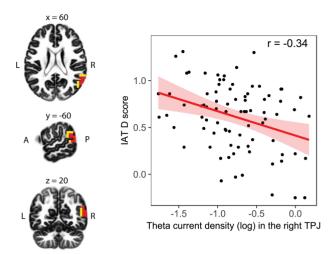


Fig. 1. Relationship between implicit intergroup bias and the baseline theta current density in the right TPJ. On the left, locations of the voxels that showed significant correlations are indicated in red (P < 0.05, corrected) and yellow (P < 0.10, corrected; displayed are MNI coordinates). On the right, the scatter plot illustrates the relationship between implicit intergroup bias (i.e. the IAT D score) and theta current density in the right TPJ (i.e. averaged current density in the theta band across all significant voxels of the right TPJ ROI, 5%; corrected, unit: A/m2), including regression line in red and confidence intervals (95%). We detected a significant negative correlation (r = -0.34, P = 0.002, 95% BCa CI [-0.14 to -0.52]) between current density within the right TPJ in the theta frequency band and implicit intergroup bias.

significantly increased the variability in implicit intergroup bias explained by personality traits alone (personality traits as predictors in Model 1: F(8,73) = 3.67, corrected $R^2 = 0.21$, P = 0.001; theta current density in right TPJ as additional predictor in Model 2: F(9,72) = 4.19, corrected $R^2 = 0.26$, P < 0.001; change in R^2 : P = 0.015; for standardized beta-values of all predictors, see Table 1).

Table 1. Predicting implicit intergroup bias by personality questionnaires and neural baseline activation. Shown are standardized beta-values and P-values of all predictors in the multiple regression analyses predicting implicit intergroup bias (IAT D-score). Model 1: personality traits only as predictors. Model 2: theta current density in right TPJ added as additional predictor

	Model 1		Model 2	
Predictor	β	P	β	P
SSIS	0.34	0.001	0.32	0.002
SDO Group-based dominance	-0.25	0.089	-0.2	0.156
SDO Opposition to equality	-0.21	0.151	-0.16	>0.20
MFQ Harm/care	0.24	0.07	0.25	0.053
MFQ Fairness/reciprocity	-0.43	0.005	-0.38	0.01
MFQ Ingroup/loyalty	-0.13	>0.20	-0.11	>0.20
MFQ Authority/respect	0.35	0.035	0.31	0.058
MFQ Purity/sanctity	-0.12	>0.020	-0.13	>0.020
Theta current density in right TPJ			-0.25	0.015

Discussion

Capitalizing on a neural trait approach, we here demonstrate that task-independent baseline current density of the right TPJ relates to individual differences in implicit intergroup bias: lower levels of theta current density in the right TPJ were associated with stronger bias. As baseline theta current density likely reflects decreased cortical activation (Scheeringa et al., 2008; Lüchinger et al., 2011; Feige et al., 2017), our findings suggest that individuals with higher activation of the right TPJ at rest exhibit stronger implicit intergroup bias. Our main finding held true for both soccer fans and political supporters, and for both men and women, demonstrating its generalizability across gender and distinct social groups. Our interpretation of the functional significance of theta current density during rest (that is, not during task execution) is based on the observation that an increase in slow wave oscillations is typically observed during lower vigilance stages and increased subjective drowsiness (e.g. Strijkstra et al., 2003). Moreover, resting EEG-fMRI studies found negative correlations between theta power and the BOLD signal in regions close to the TPJ (Scheeringa et al., 2008; Lüchinger et al., 2011; Feige et al., 2017).

Backed up by a wealth of data from metabolic neuroimaging and brain modulation studies (for reviews see Decety and Lamm, 2007; Carter and Huettel, 2013), it has been proposed that the right TPJ plays a critical role in social cognition implementing processes like self-other distinction (Uddin et al., 2006; Santiesteban et al., 2012; Sowden and Catmur, 2013; Sowden and Shah, 2014), detection of social agents (Schultz et al., 2005; Tankersley et al., 2007), perspective-taking (Ruby and Decety, 2003; Aichhorn et al., 2006) and mentalizing (Saxe and Kanwisher, 2003; Van Overwalle, 2009). In our view, categorizing in- and outgroup stimuli in the IAT does not necessitate processes such as detecting social agents, perspective-taking or mentalizing, but it does necessitate distinguishing between self-related (ingroup) and other-related (outgroup) stimuli. Consequently, the process of self-other distinction appears relevant in executing the IAT. One could thus speculate that individuals with high baseline activation of the right TPJ possess a high capacity to distinguish between the self and others, which in turn might drive their biased implicit evaluations of in- and outgroup stimuli. As our findings also revealed that right TPJ's baseline activation is capable of explaining unique variance in implicit intergroup bias, the capacity to distinguish between the self and others might represent an additional significant trait in explaining variance in implicit intergroup bias beyond relevant personality traits like in-group identification (SSIS; Wann and Branscombe, 1993), social dominance orientation (SDO; Jost and Thompson, 2000) and moral foundations (MFQ; Graham et al., 2011). Having said that, the literature on TPJ and surrounding regions in parietal cortex, also shows its role in more general cognitive processes (e.g. attention; Decety and Lamm, 2007). One could therefore speculate that higher baseline activation in the right TPJ might increase capacities in more general cognitive processes not specifically related to social cognition (Marini et al., 2018). To further elucidate this point, future studies should include a control task (see Crescentini et al., 2014), for instance an IAT unrelated to implicit intergroup bias.

Our finding that individuals with high baseline activation of the right TPJ show strong implicit intergroup bias is in line with the finding that, if the right TPJ is disrupted by means of brain modulation, behavioral intergroup bias is reduced (Baumgartner et al., 2014)—possibly mediated by a reduction in bias in the implicit evaluations that drive behavioral bias. The fact that a well-functioning right TPJ is associated with strong implicit intergroup bias suggests that biased implicit evaluations of others according to their group affiliation might fulfill a necessary function that has developed through evolution. As a social species, individuals have evolved living in different social groups (Efferson et al., 2008). In encounters with others, it was thus essential for survival to immediately determine whether someone is friend or foe (Fiske et al., 2007). These evolutionary pressures are reflected in cognitive adaptations that have developed to differentially evaluate ingroup and outgroup members. Along with the cognitive adaptations, evolution has engineered neural adaptations. Our study suggests that these neural adaptations have developed in the right TPJ.

The present study also sheds light on the role of other brain regions in driving variability in implicit intergroup bias. Interestingly, we detect no evidence in our data that the level of baseline activation of ROIs other than the right TPJ relates to bias. Thus, although task-dependent analyses have indicated that these regions are involved in intergroup bias, their task-independent neural functioning does not predispose individuals toward less or more implicit intergroup bias. For example, we found no correlation between baseline activation of the bilateral aTL and bias. It has been proposed that the aTLs are involved in processing group-related conceptual knowledge in memory (Gallate et al., 2011; Wong et al., 2012). As disrupting the aTLs reduces implicit intergroup bias, it would appear that processing conceptual knowledge in memory is involved in this phenomenon. However, this does not necessarily imply that people with a high capacity to process conceptual knowledge in memory (indicated by high baseline activation of the aTLs) demonstrate a stronger implicit intergroup bias than those with low processing capacity. Only by analyzing the relationship between neural baseline activation and bias can this issue be settled.

Finally, our findings indicate that at least partly distinct brain regions are responsible for driving variability in implicit compared to behavioral intergroup bias (for a review see Amodio, 2014). The MPFC's brain volume (another neural trait) and its connectivity with the right TPJ have been associated with variability in behavioral intergroup bias (Baumgartner et al., 2013b; Baumgartner et al., 2015), but here, we identified no association between baseline MPFC activation and implicit intergroup bias. Given its high degree of interconnectivity with diverse brain regions, the MPFC is in a position to integrate information from multiple processing streams and is thus capable of executing high-level, abstract cognitive processes (Amodio and Frith, 2006; Van Overwalle, 2009). It thus seems plausible that the MPFC is strongly involved in implementing and regulating actual behavior towards in- and outgroup members and less so in rather perceptual, low-level processes such as implicit evaluations of in- and outgroup members—which might explain why its processing capacity drives variability in behavioral, but not implicit intergroup bias.

In conclusion, the present study provides evidence that a neural trait marker—task-independent baseline activation of the right TPJ—explains individual differences in implicit intergroup bias. There was no evidence that the baseline activation of any other ROI identified in task-dependent analyses of brain regions involved in intergroup bias related to the latter, demonstrating that task-independent neural trait analyses can add a level of analysis that supplements previous task-dependent analyses (Nash et al., 2015). Overall, the present study emphasizes that neural traits are prolific targets of investigation for illuminating the sources of individual differences in implicit social cognition.

Funding

This work was supported by the European Neuroscience Network NEUREX (Grant 'The neurobiological basis of intergroup behavior in humans-insights from neuroimaging, neuropharmacological modulation, and electroencephalography' to D.K.).

Acknowledgments

We thank Y. Egenolf and M. Stein for programming the IAT in E-prime and U. Bircher, D. Chrobot, N. Gross, L. Hoffmann, E. Meissner, N. Scheurer and K. Stadler for their help in data collection and processing.

Conflict of interest. None declared.

References

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., Ladurner, G. (2006). Do visual perspective tasks need theory of mind? Neuroimage, 30(3), 1059-68.

Amodio, D.M. (2014). The neuroscience of prejudice and stereotyping. Nature Reviews Neuroscience, 15(10), 670-82.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. Nature Reviews Neuroscience, 7(4), 268-77.

Baumgartner, T., Gianotti, L.R.R., Knoch, D. (2013a). Who is honest and why: baseline activation in anterior insula predicts interindividual differences in deceptive behavior. Biological Psychology, 94(1), 192-7.

Baumgartner, T., Gotte, L., Gugler, R., Fehr, E. (2012). The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. Human Brain Mapping, 33(6), 1452-69.

Baumgartner, T., Nash, K., Hill, C., Knoch, D. (2015). Neuroanatomy of intergroup bias. A white matter microstructure study of individual differences. Neuroimage, 122, 345-54.

Baumgartner, T., Schiller, B., Hill, C., Knoch, D. (2013b). Impartiality in humans is predicted by brain structure of dorsomedial prefrontal cortex. Neuroimage, 81, 317-24.

Baumgartner, T., Schiller, B., Rieskamp, J., Gianotti, L.R.R., Knoch, D. (2014). Diminishing parochialism in intergroup conflict by disrupting the right temporo-parietal junction. Social Cognitive and Affective Neuroscience, **9**(5), 653–60.

Bergh, R., Akrami, N., Ekehammar, B. (2012). The personality underpinnings of explicit and implicit generalized prejudice. Social Psychological and Personality Science, 3(5), 614–21.

Buzsaki, G., Draguhn, A. (2004). Neuronal oscillations in cortical networks. Science, 304(5679), 1926-9.

Candidi, M., Stienen, B.M.C., Aglioti, S.M., De Gelder, B. (2015). Virtual lesion of right posterior superior temporal sulcus modulates conscious visual perception of fearful expressions in faces and bodies. Cortex, 65, 184-94.

Cannon, R.L., Baldwin, D.R., Shaw, T.L., et al. (2012). Reliability of quantitative EEG (qEEG) measures and LORETA current source density at 30 days. Neuroscience Letters, 518(1), 27-31.

Carter, R.M., Huettel, S.A. (2013). A nexus model of the temporalparietal junction. Trends in Cognitive Sciences, 17(7), 328–36.

Cattaneo, Z., Mattavelli, G., Platania, E., Papagno, C. (2011). The role of the prefrontal cortex in controlling gender-stereotypical associations: a TMS investigation. Neuroimage, 56(3), 1839-46.

Cikara, M., Van Bavel, J.J. (2014). The neuroscience of intergroup relations. An integrative review. Perspectives on Psychological Science, 9(3), 245-74.

Cohen, J. (1977). Statistical Power Analysis for the Behavioral Sciences, New York: Academic Press.

- Crescentini, C., Aglioti, S.M., Fabbro, F., Urgesi, C. (2014). Virtual lesions of the inferior parietal cortex induce fast changes of implicit religiousness/spirituality. Cortex, 54, 1-15.
- Cunningham, W.A., Nezlek, J.B., Banaji, M.R. (2004). Implicit and explicit ethnocentrism. Revisiting the ideologies of prejudice. Personality and Social Psychology Bulletin, 30(10), 1332-46.
- De Geus, E.J. (2010). From genotype to EEG endophenotype. A route for post-genomic understanding of complex psychiatric disease? Genome Medicine, 2(9), 63.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., Moors, A. (2009). Implicit measures. A normative analysis and review. Psychological Bulletin, 135(3), 347-68.
- Decety, J., Lamm, C. (2007). The role of the right temporoparietal junction in social interaction. How low-level computational processes contribute to meta-cognition. Neuroscientist, 13(6),
- Dunham, Y., Baron, A.S., Banaji, M.R. (2008). The development of implicit intergroup cognition. Trends in Cognitive Sciences, 12(7),
- Dunham, Y., Chen, E.E., Banaji, M.R. (2013). Two signatures of implicit intergroup attitudes. Developmental invariance and early enculturation. Psychological Science, 24(6), 860-8.
- Dunki, R.M., Schmid, G.B., Stassen, H.H. (2000). Intraindividual specificity and stability of human EEG. Comparing a linear vs a nonlinear approach. Methods of Information in Medicine, 39(1),
- Eagly, A.H. (1998). Attitudes and the processing of attituderelevant information. In: Adair, J.G., Bélanger, D., Dion, K.L., editors. Advances in Psychological Science: Social, Personal and Cultural Aspects, East Sussex: Psychology Press, 185-201.
- Efferson, C., Lalive, R., Fehr, E. (2008). The coevolution of cultural groups and ingroup favoritism. Science, 321(5897), 1844-9.
- Feige, B., Spiegelhalder, K., Kiemen, A., et al. (2017). Distinctive time-lagged resting-state networks revealed by simultaneous EEG-fMRI. Neuroimage, 145, 1-10.
- Fiske, S.T., Cuddy, A.J., Glick, P. (2007). Universal dimensions of social cognition. Warmth and competence. Trends in Cognitive Sciences, 11(2), 77-83.
- Friese, M., Bluemke, M., Wänke, M. (2007). Predicting voting behavior with implicit attitude measures. The 2002 german parliamentary election. Experimental Psychology, 54, 247-55.
- Gallate, J., Wong, C., Ellwood, S., Chi, R., Snyder, A. (2011). Noninvasive brain stimulation reduces prejudice scores on an implicit association test. Neuropsychology, 25(2), 185-92.
- Gianotti, L.R.R., Figner, B., Ebstein, R.P., Knoch, D. (2012). Why some people discount more than others. Baseline activation in the dorsal PFC mediates the link between COMT genotype and impatient choice. Frontiers in Neuroscience, 6, 54.
- Gianotti, L.R.R., Lobmaier, J.S., Calluso, C., Dahinden, F.M., Knoch, D. (2018a). Theta resting EEG in TPJ/pSTS is associated with individual differences in the feeling of being looked at. Social Cognitive and Affective Neuroscience, 13(2), 216–23.
- Gianotti, L.R.R., Nash, K., Baumgartner, T., Dahinden, F.M., Knoch, D. (2018b). Neural signatures of different behavioral types in fairness norm compliance. Scientific Reports, 8, 10513.
- Graham, J., Nosek, B.A., Haidt, J., Iyer, R., Koleva, S., Ditto, P.H. (2011). Mapping the moral domain. Journal of Personality and Social Psychology, 101(2), 366-85.
- Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition. The implicit association test. Journal of Personality and Social Psychology, 74(6), 1464-80.
- Greenwald, A.G., Nosek, B.A., Banaji, M.R. (2003). Understanding and using the implicit association test. I. An improved scor-

- ing algorithm. Journal of Personality and Social Psychology, 85(2),
- Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L., Banaji, M.R. (2009). Understanding and using the implicit association test. III. Meta-analysis of predictive validity. Journal of Personality and Social Psychology, **97**(1), 17–41.
- Hahn, T., Notebaert, K., Anderl, C., et al. (2015a). Reliance on functional resting-state network for stable task control predicts behavioral tendency for cooperation. Neuroimage, 118,
- Hahn, T., Notebaert, K., Anderl, C., Teckentrup, V., Kassecker, A., Windmann, S. (2015b). How to trust a perfect stranger. Predicting initial trust behavior from resting-state brain-electrical connectivity. Social Cognitive and Affective Neuroscience, 10(6),
- Harmon-Jones, E., Gable, P.A., Peterson, C.K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena. A review and update. Biological Psychology, 84(3),
- Hayes, A.F. (2013). Introduction to Mediation, Moderation, and Conditional Process Analysis, New York: The Guilford Press.
- Hein, G., Silani, G., Preuschoff, K., Batson, C.D., Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. Neuron, 68(1),
- Hewstone, M., Rubin, M., Willis, H. (2002). Intergroup bias. Annual Review of Psychology, 53, 575-604.
- Jost, J.T., Thompson, E.P. (2000). Group-based dominance and opposition to equality as independent predictors of selfesteem, ethnocentrism, and social policy attitudes among African Americans and European Americans. Journal of Experimental Social Psychology, 36(3), 209–32.
- Korb, S., Malsert, J., Rochas, V., et al. (2015). Gender differences in the neural network of facial mimicry of smiles—An rTMS study. Cortex, 70, 101-14.
- Knoch, D., Gianotti, L.R.R., Baumgartner, T., Fehr, E. (2010). A neural marker of costly punishment behavior. Psychological Science, **21**(3), 337–42.
- Kubicki, S., Herrmann, W.M., Fichte, K., Freund, G. (1979). Reflections on the topics. EEG frequency bands and regulation of vigilance. Pharmakopsychiatrie Und Neuropsychopharmakologie, **12**(2), 237–45.
- Kubota, J.T., Banaji, M.R., Phelps, E.A. (2012). The neuroscience of race. Nature Neuroscience, 15, 940–8.
- Kubota, J.T., Li, J., Bar-David, E., Banaji, M.R., Phelps, E.A. (2013). The price of racial bias. intergroup negotiations in the ultimatum game. Psychological Science, 24(12), 2498-504.
- Lancaster, J.L., Woldorff, M.G., Parsons, L.M., et al. (2000). Automated Talairach Atlas labels for functional brain mapping. Human Brain Mapping, 10(3), 120-31.
- Lopes da Silva, F. (2013). EEG and MEG. Relevance to neuroscience. Neuron, 80(5), 1112-28.
- Lüchinger, R., Michels, L., Martin, E., Brandeis, D. (2011). EEG-BOLD correlations during (post-)adolescent brain maturation. Neuroimage, **56**(3), 1493–505.
- Marini, M., Banaji, M.R., Pascual-Leone, A. (2018). Studying implicit social cognition with noninvasive brain stimulation. Trends in Cognitive Sciences, 22(11), 1050-66.
- Mazziotta, J., Toga, A., Evans, A., et al. (2001). A probabilistic atlas and reference system for the human brain. International Consortium for Brain Mapping (ICBM). Philosophical Transactions of the Royal Society B: Biological Sciences, 356(1412), 1293-322.

- Meng, X., Rosenthal, R., Rubin, D.B. (1992). Comparing correlated correlation coefficients. Psychological Bulletin, 111(1), 172-5.
- Mierke, J., Klauer, K.C. (2001). Implicit association measurement with the IAT: evidence for effects of executive control processes. Experimental Psychology, 48, 107-22.
- Mobascher, A., Brinkmeyer, J., Warbrick, T., et al. (2009a). Fluctuations in electrodermal activity reveal variations in single trial brain responses to painful laser stimuli—A fMRI/EEG study. Neuroimage, 44(3), 1081-92.
- Mobascher, A., Brinkmeyer, J., Warbrick, T., et al. (2009b). Laser-evoked potential P2 single-trial amplitudes covary with the fMRI BOLD response in the medial pain system and interconnected subcortical structures. Neuroimage, 45(3), 917-26.
- Molenberghs, P. (2013). The neuroscience of in-group bias. Neuroscience & Biobehavioral Reviews, 37(8), 1530-6.
- Näpflin, M., Wildi, M., Sarnthein, J. (2007). Test-retest reliability of resting EEG spectra validates a statistical signature of persons. Clinical Neurophysiology, 118(11), 2519-24.
- Nash, K., Gianotti, L.R.R., Knoch, D. (2015). A neural trait approach to exploring individual differences in social preferences. Frontiers in Behavioral Neuroscience, 8, 458.
- Nichols, T.E., Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging. A primer with examples. Human Brain Mapping, 15(1), 1-25.
- Nosek, B.A., Hawkins, C.B., Frazier, R.S. (2011). Implicit social cognition: from measures to mechanisms. Trends in Cognitive Sciences, 15(4), 152-9.
- Nosek, B.A., Sriram, N. (2007). Faulty assumptions. A comment on Blanton, Jaccard, Gonzales, and Christie (2006). Journal of Experimental Social Psychology, 43(3), 393-8.
- Nuwer, M.R., Comi, G., Emerson, R., et al. (1998). IFCN standards for digital recording of clinical EEG. International Federation of Clinical Neurophysiology. Electroencephalography and Clinical Neurophysiology, 106(3), 259-61.
- Pascual-Marqui, R.D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA). Technical details. Methods and Findings in Experimental and Clinical Pharmacology, 24(Suppl D), 5-12.
- Pratto, F., Shih, M. (2000). Social dominance orientation and group context in implicit group prejudice. Psychological Science, 11(6), 515-8.
- Raccuia, L. (2016). Single-target implicit association tests (ST-IAT) predict voting behavior of decided and undecided voters in swiss referendums. PLoS One, 11(10), e0163872 10.
- Rowatt, W.C., Franklin, L.M., Cotton, M. (2005). Patterns and personality correlates of implicit and explicit attitudes toward Christians and Muslims. Journal for the Scientific Study of Religion, 44(1), 29-43.
- Ruby, P., Decety, J. (2003). What you believe versus what you think they believe. A neuroimaging study of conceptual perspective-taking. European Journal of Neuroscience, 17(11),
- Rullmann, M., Anwander, A., Dannhauer, M., Warfield, S.K., Duffy, F.H., Wolters, C.H. (2009). EEG source analysis of epileptiform activity using a 1 mm anisotropic hexahedra finite element head model. Neuroimage, 44(2), 399-410.
- Santiesteban, I., Banissy, M.J., Catmur, C., Bird, G. (2012). Enhancing social ability by stimulating right temporoparietal junction. Current Biology, 22(23), 2274-7.
- Saxe, R., Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". Neuroimage, 19(4), 1835-42.

- Scheeringa, R., Bastiaansen, M.C., Petersson, K.M., Oostenveld, R., Norris, D.G., Hagoort, P. (2008). Frontal theta EEG activity correlates negatively with the default mode network in resting state. International Journal of Psychophysiology, 67(3), 242-51.
- Schiller, B., Baumgartner, T., Knoch, D. (2014a). Intergroup bias in third-party punishment stems from both ingroup favoritism and outgroup discrimination. Evolution and Human Behavior, 35(3), 169-75.
- Schiller, B., Gianotti, L.R.R., Baumgartner, T., Nash, K., Koenig, T., Knoch, D. (2016). Clocking the social mind by identifying mental processes in the IAT with electrical neuroimaging. Proceedings of the National Academy of Sciences of the United States of America, 113(10), 2786-91.
- Schiller, B., Gianotti, L.R.R., Nash, K., Knoch, D. (2014b). Individual differences in inhibitory control—relationship between baseline activation in lateral PFC and an electrophysiological index of response inhibition. Cerebral Cortex, 24(9), 2430-5.
- Schultz, J., Friston, K.J., O'Doherty, J., Wolpert, D.M., Frith, C.D. (2005). Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. Neuron, 45(4),
- Smit, D.J., Posthuma, D., Boomsma, D.I., Geus, E.J. (2005). Heritability of background EEG across the power spectrum. Psychophysiology, 42(6), 691-7.
- Smith, I.H., Aquino, K., Koleva, S., Graham, J. (2014). The moral ties that bind . . . Even to out-groups. the interactive effect of moral identity and the binding moral foundations. Psychological Science. 25(8), 1554-62.
- Sowden, S., Catmur, C. (2013). The role of the right temporoparietal junction in the control of imitation. Cerebral Cortex, 25(4), 1107-13.
- Sowden, S., Shah, P. (2014). Self-other control: a candidate mechanism for social cognitive function. Frontiers in Human Neuroscience, 8, 789.
- Strijkstra, A.M., Beersma, D.G., Drayer, B., Halbesma, N., Daan, S. (2003). Subjective sleepiness correlates negatively with global alpha (8-12 Hz) and positively with central frontal theta (4-8 Hz) frequencies in the human resting awake electroencephalogram. Neuroscience Letters, 340(1), 17–20.
- Strombach, T., Weber, B., Hangebrauk, Z., et al. (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. Proceedings of the National Academy of Sciences of the United States of America, 112(5), 1619-24.
- Tankersley, D., Stowe, C.J., Huettel, S.A. (2007). Altruism is associated with an increased neural response to agency. Nature Neuroscience, **10**(2), 150–1.
- Uddin, L.Q., Molnar-Szakacs, I., Zaidel, E., Iacoboni, M. (2006). rTMS to the right inferior parietal lobule disrupts self-other discrimination. Social Cognitive and Affective Neuroscience, 1(1), 65-71.
- Van Diessen, E., Numan, T., van Dellen, E., et al. (2015). Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. Clinical Neurophysiology, 126(8), 1468-81.
- Van Overwalle, F. (2009). Social cognition and the brain: a metaanalysis. Human Brain Mapping, 30(3), 829-58.
- Wann, D.L., Branscombe, N.R. (1993). Sports fans-measuring degree of identification with their team. International Journal of Sport Psychology, 24(1), 1–17.
- Wong, C.L., Harris, J.A., Gallate, J.E. (2012). Evidence for a social function of the anterior temporal lobes: low-frequency rTMS reduces implicit gender stereotypes. Social Neuroscience, 7(1), 90-104.

Wozniak-Kwasniewska, A., Szekely, D., Aussedat, P., Bougerol, T., David, O. (2013). Changes of oscillatory brain activity induced by repetitive transcranial magnetic stimulation of the left dorsolateral prefrontal cortex in healthy subjects. Neuroimage, 88, 91–9.

Zietsch, B.P., Hansen, J.L., Hansell, N.K., Geffen, G.M., Martin, N.G., Wright, M.J. (2007). Common and specific genetic influences on EEG power bands delta, theta, alpha, and beta. Biological Psychology, 75(2), 154-64.