



Early neural markers of implicit attitudes: N170 modulated by intergroup and evaluative contexts in IAT

Agustín Ibáñez^{1,2,3,4*}, Ezequiel Gleichgerrcht^{1,4}, Esteban Hurtado^{3,5}, Ramiro González^{3,5}, Andrés Haye⁵ and Facundo F. Manes^{1,4}

¹ Laboratory of Experimental Psychology and Neurosciences, Institute of Cognitive Neurology (INECO), Buenos Aires, Argentina

² National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina

³ Laboratory of Cognitive Neuroscience, Universidad Diego Portales, Santiago, Chile

⁴ Institute of Neuroscience, Favaloro University, Buenos Aires, Argentina

⁵ School of Psychology, Pontificia Universidad Católica de Chile, Santiago, Chile

Edited by:

Chris Rorden, Georgia Institute of Technology, USA

Reviewed by:

Lutz Jäncke, University of Zurich, Switzerland

Bruce D. Bartholow, University of Missouri, USA

*Correspondence:

Agustín Ibáñez, Laboratory of Experimental Psychology and Neurosciences, Institute of Cognitive Neurology and National Scientific and Technical Research Council, Castex 3293 (CP 1425) Buenos Aires, Argentina.

e-mail: aibanez@neurologiacognitiva.org

The Implicit Association Test (IAT) is the most popular measure to evaluate implicit attitudes. Nevertheless, its neural correlates are not yet fully understood. We examined event related potentials (ERPs) in response to face- and word processing while indigenous and non-indigenous participants performed an IAT displaying faces (ingroup and outgroup members) and words (positive and negative valence) as targets of category judgments. The N170 component was modulated by valence of words and by ingroup/outgroup face categorization. Contextual effects (face–words implicitly associated in the task) had an influence on the N170 amplitude modulation. On the one hand, in face categorization, right N170 showed differences according to the association between social categories of faces and affective valence of words. On the other, in word categorization, left N170 presented a similar modulation when the task implied a negative-valence associated with ingroup faces. Only indigenous participants showed a significant IAT effect and N170 differences. Our results demonstrate an early ERP blending of stimuli processing with both intergroup and evaluative contexts, suggesting an integration of contextual information related to intergroup attitudes during the early stages of word and face processing. To our knowledge, this is the first report of early ERPs during an ethnicity IAT, opening a new branch of exchange between social neuroscience and social psychology of attitudes.

Keywords: N170, VPP, dual valence task, attitudes, implicit association test, contextual blending, intergroup context, indigenous participants

INTRODUCTION

Ingroup favoritism has often been regarded as the basis for positive membership feelings, safety, and familiarity. At the same time, however, it is known to be one of the forces that drive discrimination, religious battles, and war. Social psychology has a long tradition of studying the cognitive and affective bases of prejudice and social bias (Allport, 1954; Devine, 1989; Glick and Fiske, 2001; Gawronski et al., 2008). In fact, contemporary research on prejudice claims that an individual's attitude toward a member of an outgroup essentially stems from automatic associative processes that may explain biased behavior (Bargh, 1997; Fazio, 2001; Nosek and Banaji, 2001). According to the leading view, these associative processes need neither effort nor cognitive attention, and they may even occur in the absence of conscious awareness (for a discussion, see Gawronski et al., 2006). For this reason, implicit measurement techniques are particularly useful in the assessment of bias in attitudes toward objects associated with ingroups vs. outgroups. Indeed, such techniques are assumed to reflect introspectively unidentified or inappropriately identified traces of past experiences (Fazio and Olson, 2003; Greenwald et al., 2003).

One of the most popular implicit attitude measures employed in social psychology is the Implicit Association Test (IAT; Greenwald et al., 1998). The IAT is a simple task in which participants are asked to make category and/or evaluative judgments as quickly

as possible, thus measuring the strength of associations between two categories. Widely used in social psychology (De Houwer et al., 2009), the IAT is implicit because participants are not asked about the associations *per se*. Rather, their reaction times to these associations are recorded. Extensive research has demonstrated the usefulness of the IAT in the assessment of social bias by considering social categories (e.g., pictures of Black and White people) and valence categories (i.e., ranging from positive to negative). As well, IATs are valuable in determining whether experimental blocks that require “compatible” associations (e.g., in prejudice: White person/positive or Black person/negative) show increased or decreased reaction times relative to “incompatible” associations (e.g., White person/negative or Black person/positive). Note that the IAT effect is triggered by the *contextual association* between faces and words. For example, a block of IAT trials considered to be “compatible” with prejudice implies a white participant responding to pleasant words and white faces with one key, and to unpleasant words and black faces with another key, in order to assess delays in response times. In “incompatible” blocks of trials, response pairings are inverted (i.e., black–positive, white–negative). It is argued that prejudiced participants have strong “compatible” associations, and that “incompatible” associations would be weaker, as revealed by the response latency differences that may exist between these two possible scenarios. In this sense, the IAT is meant as a measure of

strength of automatic association (Greenwald et al., 2003), because by assuming that the categorization of stimuli according to automatic associations is easier and hence faster, one can interpret relatively shorter latencies in “compatible” trials to be reflective of stronger racial bias.

Although this is one of the most widely used methods among implicit paradigms of attitudes research, there is some extent of controversy concerning its intrinsic significance as a measure of attitude. This is mainly because of the theoretical assumptions about the cognitive and behavioral processes involved in the IAT task (e.g., Arkes and Tetlock, 2004; Hamann and Canli, 2004; Rothermund and Wentura, 2004; Conrey et al., 2005; Kinoshita and Peek-O’Leary, 2005, 2006; Blanton et al., 2009). Besides the specific criticisms toward particular implicit attitude measurement techniques, the broader question targeting the underlying processes in attitudes has acquired increased importance. The present study seeks to tackle part of this general issue. For example, in social psychology, the discussion concerning implicit attitudes measures has raised interest in the distinction and possible relationships between implicit and explicit (conscious, introspective, overt) processes, as involved in prejudice toward minorities or lower-status social groups. Numerous findings on implicit and explicit attitudes have been reported on Black and White individuals (Dasgupta et al., 2000), Asians (Dunham et al., 2006), Germans and Turks (Bohner et al., 2008), New Zealand (Sibley and Liu, 2007), and Chilean indigenous people (Brown et al., 2008; Hurtado et al., 2009; Ibáñez et al., 2009; Haye et al., 2010); among others. The present study deals with implicit biases involving Chilean indigenous people, specifically, the Mapuche ethnic group. The Mapuche people are Chile’s largest indigenous group, with a long history of social deprivation (see Bengoa, 2000; Dillehay, 2007) and typically associated with negative beliefs about them by the Chilean non-indigenous society. To them, Mapuche people are stereotyped as violent, rude, lazy, and unintelligent (Saiz and Williams, 1991).

Different from these previous studies in social psychology, our previous work engages the social neuroscience approach to cognitive context-dependent effects (Ibáñez et al., 2006, 2009, 2010a,b,c; Cornejo et al., 2009; Hurtado et al., 2009; Aravena et al., 2010; Riveros et al., 2010; San Martín et al., 2010), and focuses on the possible relationships between behavioral levels of description, as addressed by the IAT, and early electrophysiological phenomena. In particular, we studied early electrophysiological changes associated with performing an ethnic IAT. Only one previous report has assessed late event related potential (ERP) components in response to an ethnic IAT. Hurtado et al. (2009) demonstrated that, during completion of an IAT, indigenous participants with greater behavioral ingroup bias displayed a frontal late positive potential (LPP) that was modulated in terms of complex contextual associations featuring both valence and ethnic group. In the present study, we report the early electrophysiological responses to the previous reported LPP effects. In doing so, we aim to understand automatic brain measures concerning basic steps of attitude processes, as the gateway to a multilevel approach of this issue.

Ethnic cues such as facial features evolve together with other cultural traits, and are used as signals of intergroup relative position (Efferson et al., 2008). Facial features may have a special importance since face perception is the most highly developed visual skill

(Gobbini and Haxby, 2006). Most likely, discrimination of ingroup/outgroup members involves a process that is both simultaneously top-down and bottom-up: within a few milliseconds, extremely subtle facial cues can be discriminated, which may allow for the distinction between ingroup and outgroup members. Furthermore, processes involving contextual associations (for example, in our paradigm, a face that is contextually associated with a negative- or positive-valence word) may intervene in this perception and its subsequent evaluation. However, little is known about the early integration of intergroup (e.g., racial) cues and valence cues. Here, we assessed the early neural processing of intergroup and evaluative contextual information regarding facial and verbal stimuli, embedded in an ethnic IAT, as described above and further detailed in the Section “Materials and Methods.” To our knowledge, no previous study of IAT with early electrophysiological correlates has been previously reported. Although there are some recent studies on early ERP effects on same/other-race faces in an N170 task that has been off-line correlated with IAT (e.g., He et al., 2009), this is the first report of N170 effects directly elicited by the IAT.

There are some major questions that must be tackled, namely: what early physiological changes can be identified during the performance of an IAT trial as a function of the intergroup position of the target relative to the participant’s social membership? Will an early electrophysiological modulation by membership (faces) and valence (words), as well as the possible contextual association effects be observed? Or, will the modulation be based on the compatibility effects? Will an inversion of neural markers elicited by the valence and membership of stimuli in dependence of the participant’s relative position be found? Because of the complex interaction between these variables, we are only able to state one general hypothesis in advance: If ingroup/outgroup distinction plays a role in automatic evaluations – as suggested by the literature – then the intergroup relative position as well as the contextual association effects should make a difference in early stimulus processing, thus systematically modulating neural activity just a few milliseconds after exposure.

EARLY ELECTROPHYSIOLOGICAL PHENOMENA

Event related potentials constitute a technique that allows to measure early brain activity in response to the categorization of a face or a word stimulus in challenging reaction-time tasks, such as the IAT. ERPs, which are measured using electroencephalography (EEG), estimate the ongoing electrophysiological changes resulting from the synchronous activation of several neural subpopulations that occur in response to sensory, motor or cognitive events (Hillyard and Picton, 1987).

N170 and Vertex Positive Potential (VPP) are two components expected to be modulated by face stimuli, such as those employed in IATs. ERP studies have identified that between 130 and 170 ms, early cerebral modulation occurs, which is caused by stimuli from images of faces and other objects in the occipito-temporal scalp (N170; Bentin et al., 1996) which is temporally simultaneous with – and functionally equivalent to – modulation in the fronto-central scalp regions (VPP; Joyce and Rossion, 2005). It has been suggested that this N170/VPP complex may be generated by the inferior temporal gyrus and fusiform gyrus (Horowitz et al., 2004), both of which are associated with facial processing. Accordingly, the N170/VPP complex has shown to be sensitive to face categorization

(Linkenkaer-Hansen et al., 1998; Rossion et al., 1999, 2000, 2002; Rebai et al., 2001; Itier and Taylor, 2002), while stimuli which are neither faces nor words do not typically produce the N170/VPP effect (Rossion et al., 2000, 2002; Rebai et al., 2001). Whether N170 and VPP are part of the same neural dipole is still a matter of debate, especially because of their temporal overlap. Because N170 seems to be more influenced by emotional and contextual clues than VPP, we focused more on this former component.

Studies that use face inversion suggest that N170 reflects an early structural/holistic process of perceptual coding, as face inversion seems to disrupt facial feature integration (Sagiv and Bentin, 2001; Rossion et al., 2002; Honda et al., 2007; Ishizu et al., 2008; Jemel et al., 2009). N170 has been described as lateralized in the right hemisphere for faces, and in the left hemisphere for words, and bilaterally for object categorization (Brandeis et al., 1995; Rossion et al., 2003). Previous studies of attentional focus on the processing of racial cues from faces have shown N170 modulation (e.g., Ito and Urland, 2005). However, other studies using “other-race” face classification paradigms failed to show such N170 modulation of race (Caldara et al., 2003, 2004). Besides the race effects associated with N170, some reports have suggested the modulation of blended attributes (e.g., facial expression) with other features of stimuli (e.g., affective valence). Several studies have shown N170 modulation based on faces and emotions (Pizzagalli et al., 2002; Batty and Taylor, 2003; Ashley et al., 2004; Blau et al., 2007; Krombholz et al., 2007). Some results suggest that facial processing involves multiple mechanisms in which structural (i.e., orientation) and semantic (i.e., familiarity) factors begin to interact in early stages with different time courses (Marzi and Viggiano, 2007). Galli et al. (2006) demonstrated that early stages of structural analysis may be influenced by the presence of meaningful context. In brief, in considering N170 and VPP as distinguishable components, the former, and partially the latter, are potential neural markers of early (N170) and selective face processing (and less influenced by word processing), modulated by emotional content, affective valence, and contextual clues. Based on these previous reports, we expected an N170 modulation of relative intergroup position of face targets, and valence of word targets, as well as contextual effects between faces and words in the IAT. As stated above, previous ERP studies suggest an early sensitivity of brain face processing to ethnic, emotional, and contextual clues. Although there are prior reports of early effects of race and evaluative meaning in the literature, to the best of our knowledge, there is no report of early (N170/VPP) contextual association of ingroup/outgroup stimuli, valence of words and ethnicity of participants using an IAT. Although some reports had shown an earlier than N170 emotional effect (e.g., P1 component: Pourtois et al., 2004), we did not observe any emotional/racial effect earlier than the N170 window (see P1 Effects in Supplementary Material). Consequently, we focused on the structural stage of facial processing indexed by the N170/VPP.

In this study, we recorded ERPs while indigenous (Mapuche) and non-indigenous participants performed an IAT with faces (ingroup and outgroup) and words (positive and negative valence). By including both participant groups (indigenous and non-indigenous), comparable in age and gender, we were able to compare the intergroup relative position of faces and its contextual influence on early processing for both positions. We hypothesized that

the intergroup and evaluative contexts involved in the IAT task would have an effect during early stages of processing (N170/VPP). Positive-valence information should increase the N170 amplitude (and potentially, the VPP as well). Outgroup information may also increase the amplitude of N170. Moreover, we expected to find a different ERP pattern as a function of the association implied by the IAT trial between faces and words. For example, an outgroup stimulus (indigenous face presented to a non-indigenous participant) should elicit different N170 amplitude when paired with a positive-valence word (incongruent contextual association) than with a negative word (congruent contextual association). Those amplitude effects should be the opposite (inverted) between groups (indigenous vs. non-indigenous participants). Additionally, if the behavioral measures of ingroup bias (IAT score) in one group are stronger, an increased early modulation of ERPs by intergroup and evaluative contexts should be observed.

MATERIALS AND METHODS

PARTICIPANTS

Thirty-six Chilean subjects, between 18 and 40 years old, participated in the study. Half were of indigenous origin (mean age 26.2 years, $SD = 7.13$) and half were non-indigenous (mean age 25.7 years, $SD = 5.70$). Indigenous origin was determined by participants identifying themselves with the Mapuche ethnic group and by having a Mapuche last name. The non-indigenous participants did not identify themselves with any native Chilean ethnic group and did not come from indigenous families. Age was comparable between the groups ($t = 0.44$, $df = 34$, $p = 0.65$). Of the 18 indigenous participants, 11 were women (61%) and in the non-indigenous group, 7 were women (38%; $t = 1.33$, $df = 34$, $p = 0.19$). All participants were right-handed (assessed with the Edinburgh Handedness Inventory) without any visual deficits or psychopathology, as assessed during the initial interview. Regarding educational level, we did not find significant differences between the groups [Mapuche: $M: 10.33$ years, $SD: 2.81$; Non-Indigenous: $M: 11.42$ years, $SD: 3.21$; $F(1, 32) = 0.22$, $p = 0.74$]. The non-indigenous participants were students, while the indigenous participants were contacted through an indigenous community in Santiago. All signed an informed consent accepting their participation in the study and protection of their identity in accordance with the Helsinki Declaration.

PROCEDURE

Initially, the participants were informed that the study assessed recognition and classification processes regarding people and words. For the IAT, after presenting the instructions, participants sat in front of a computer with the electrodes placed on their heads, and completed the task with the stimuli being displayed on the computer screen. Behavioral responses were generated by pressing with each forefinger any of two keys (1 vs. 2) on a response pad (see Hurtado et al., 2009 for a detailed description of the experimental procedure). Once the experiment had finished, participants were thanked and the research goal was thoroughly explained.

IAT TASK

Several steps for validation of verbal and pictorial stimuli for IAT procedure were performed (see IAT Stimuli Validation in Supplementary Material; or Hurtado et al., 2009). The task was

run using E-prime software. All stimuli were in Spanish and presented in different trial blocks. Each block began with a brief explanation of which classification category, or category combination, was assigned to each response key. Subsequently, the trials were presented one by one, until the specific number for each block was displayed. The practice blocks involved approximately 28 stimuli, and the test blocks 100 stimuli. In order to assess the correct detection of stimuli faces were displayed for 100 ms and words for 300 ms (Hurtado et al., 2009). Incorrect responses were indicated with an “X” in the center of the screen following the response. The stimuli were both horizontally and vertically centered on the screen. For each block, the categories to be used for the classification of the target stimuli were displayed on the top left and right-hand corners. Category names were “Indigenous,” “Non-indigenous,” “Pleasant,” and “Unpleasant.”

A general IAT design was applied (Greenwald et al., 1998), with minor modifications. In block 1, participants categorized faces either as indigenous or non-indigenous. In block 2, they were asked to categorize words either as pleasant or unpleasant. For block 3, intergroup and evaluative categories of blocks 1 and 2 were pooled together, respectively, in a single task of combined classification, such that stimuli were categorized as Indigenous–Unpleasant and Non-indigenous–Pleasant. Hence, this block was of a “compatible with prejudice toward the minority” type. This discrimination task was carried out in practice block 3 and repeated with more stimuli in test block 4. In block 5, words were categorized, but this time categories were in opposite top corners compared to previous blocks. In block 6, faces were also categorized. After applying inverted category–key conditions in blocks 5 and 6, the combined classification tasks followed in blocks 7 and 8, specifically categorizing the four types of stimuli as either Non-indigenous–Unpleasant or Indigenous–Pleasant, thus being of the “incompatible” type. Block 7 was a practice block, whereas block 8 was the corresponding test block (Greenwald et al., 1998). As well, in order to counterbalance the assignment of hands to categories, the series of 8 blocks was carried out twice, thus producing a total of 16 blocks. Each trial began with a fixation cross presented before the stimuli (1000 ms). A blank screen after the presentation of stimuli was presented until the response. For wrong responses, a negative feedback (“X”) was presented for 700 ms. An inter-trial of 1000 ms was included.

ELECTROPHYSIOLOGICAL RECORDINGS

Signals were recorded on-line using a GES300, 129-channel system with HydroCel Sensors from Electrical Geodesic Inc with a DC coupling Amplifier, 24-bit A/D converter, 200 M_{in} input impedance, 0.7 μV RMS/1.4 μV pp noise; and the NetStation™ software. Analog filters were 0.1 and 100 Hz. A digital band pass filter between 0.5 and 30 Hz was applied off-line to remove unwanted frequency components. Signals were sampled at 500 Hz. During recordings, the reference was set by default to vertex but then was re-referenced off-line to average electrodes. Two bipolar derivations were designed to monitor vertical and horizontal ocular movements (EOG). One-second stimulus-locked epochs were selected from the continuous data, beginning 200 ms prior to stimulus onset. Baseline was removed at –200 to 0 ms before stimulus onset. All epochs with eye movement contamination were removed from further analysis,

using automatic (Gratton, Coles, and Donchin method for removing eye-blink artifacts) and visual procedures. Artifact-free epochs were averaged to obtain the ERPs. The EEGLAB Matlab toolbox and T-Besp software (<http://neuro.udp.cl/software>) were used for EEG off-line processing and analysis.

Event related potential waveforms were averaged separately for each of the following eight experimental conditions: Indigenous picture in an Indigenous–positive context; Non-indigenous picture in a Non-indigenous–positive context; Indigenous picture in an Indigenous–negative context; Non-indigenous picture in a Non-indigenous–negative context; Positive word in an Indigenous–positive context; Positive word in a Non-indigenous–positive context; Negative word in an Indigenous–negative context; Negative word in a Non-indigenous–negative context.

STATISTICAL ANALYSES

Behavioral measures (IAT)

An implicit racial bias rate was calculated for each subject based on reaction times obtained from the IAT test. This numeric value is a modification of Greenwald et al. (2003) which provides an indication of the tendency observed in the reaction-time difference between “compatible” and “incompatible” trials with ingroup bias. Reaction times obtained from the test were used to determine an IAT score for each participant such that it would provide an index of the tendency observed in the reaction-time difference between compatible and incompatible tasks. Greenwald et al. (2003) proposed a method to calculate this rate for social categorization IATs, which contemplates the elimination of extreme reaction times, special arrangement of wrong responses and standardization of resulting reaction times based on the standard of each participant (see Extraction of IAT Measures: Algorithm Proposed by Greenwald et al., 2003 in Supplementary Material or Hurtado et al., 2009; for more details about this algorithm). A participant’s rate was obtained by measuring the reaction-time difference between compatible and incompatible blocks, standardized according to the standard deviation. Therefore, the measurement procedure was evaluated for bias using a scale that allows for comparison between different participants. This procedure led to an individual score for each participant with an expected value close to zero for those who did not show any bias in the test, and positive values associated with the detection of bias in favor of the compatible task. The relevant factors in the behavioral IAT were: participants (indigenous/non-indigenous), task (compatible/incompatible) and stimulus type (faces/words). We include this last comparison since previous reports had sown a differential effect of faces vs. words (see Foroni and Bel-Bahar, 2009). Therefore, the results were grouped in three measurement sets. Welch’s *t*-tests were used on two samples and two tails. The resulting four *p*-values were adjusted by the Holm–Bonferroni correction, in order to control for the possibility of a type I error. Therefore, a criterion was used to determine which experimental groups presented racial bias based on the IAT test and what type of stimulus generated said effect. Analysis of variance (ANOVA) was applied to explore the tendencies in the results, using a mixed model that had each subject as a random effect linked to the three factors. For significant effects ($p = 0.05$), averages and contrasts were calculated with Tukey’s *post hoc* tests.

ERPs

Regions of interest (ROI) were used to represent and analyze the scalp topography of the ERP components as recommended for dense arrays, in order to avoid loss of statistical power (Oken and Chiappa, 1986). The ERP analysis was carried out based on the ROIs chosen by visually checking each component. Each ROI comprised six adjacent electrode sites centered around the maxima on Cz, P8 and P7 (Rossion and Jacques, 2008):

1. VPP (maximum in Cz, EGIs electrodes 129 [Cz], 7 106 13 6 112)
2. Left N170: (maximum in P7, EGIs electrodes 59 [P7] 58 65 64 69 68)
3. Right N170 (maximum in P8, EGIs electrodes 91 [P8] 96 90 95 99 94)

Although the figures show the ERPs grand averages from each group, all statistical calculations were conducted using each participant's individual data. ERP amplitudes were quantified as the mean average from negative (for N170) or positive (for VPP) deflection occurring within a 140–190 ms temporal window (N170/VPP).

The ERPs analysis was performed considering faces and words separately. Consequently, for face targets, the factors included were intergroup relative position (IP hereafter: ingroup vs. outgroup face) and contextual association of words (positive vs. negative). For word targets, the analogous structure was word target valence (positive vs. negative) and contextual association of faces (ingroup vs. outgroup). In that way, main effects of target and context specifications as well as potential interactions were detected. The inclusion of the *contextual association* factor was considered in order to detect the contextual task-related associations. For face targets, this factor describes the word's valence associated context. For word targets, it refers to face (ingroup or outgroup) associations.

Finally, the ROI factor (left vs. right for N170) was included when necessary. A mixed ANOVA of repeated measures with three within-subject factors was performed for faces (IP, contextual association, and ROI) and for words (valence, contextual association, and ROI). For VPP analysis no different ROIs were tested. A between-subject factor with participant's group (indigenous vs. non-indigenous) was also included. Univariate comparisons were carried out whenever necessary. The average number of artifact-free trials across conditions was 84.3% for non-indigenous and 81.4% for indigenous groups.

RESULTS

BEHAVIORAL RESULTS

IAT effect

As previously reported in Hurtado et al. (2009), the indigenous group obtained an average score of 0.46 (SD = 0.57). As expected, this value was significantly higher than zero ($t = 3.43$, $df = 17$, $p < 0.01$), which means that this group took more time to respond to the compatible than the incompatible task. The opposite was observed in the non-indigenous group, which obtained an average of 0.07 (SD = 0.73); in this case, the value was not significantly different from zero ($t = -0.43$, $df = 17$, $p = 0.66$). As for word stimuli only, the indigenous group obtained an average score of 0.71 (SD = 0.68), significantly different from zero ($t = 4.48$, $df = 17$,

$p = 0.0003$); the non-indigenous group obtained an average of 0.01 (SD = 0.84) which was not significantly different from zero ($t = 0.08$, $df = 17$, $p = 0.93$). In terms of face stimuli only, indigenous participants obtained an average score of 0.38 (SD = 0.85), whereas the non-indigenous group obtained an average of 0.17 (SD = 0.70). These face scores were neither significant from zero for the indigenous ($t = 1.89$, $df = 17$, $p = 0.07$) nor for the non-indigenous ($t = -1.07$, $df = 17$, $p = 0.29$). After Holm adjustment of p -values, only the indigenous group's IAT score for words ($p = 0.002$) and the general score for the same group ($p = 0.01$) remained significantly different from zero. These measures revealed that the test applied detected racial bias within the indigenous group (General and Word IAT conditions) but not among the members of the non-indigenous group.

ERPs

Visual inspection of data for both groups collapsed (see **Figure 1**) suggested a N170/VPP effect of modulation as a function of IP, and valence of words. In order to examine this modulation in detail, each component was analyzed separately.

VPP

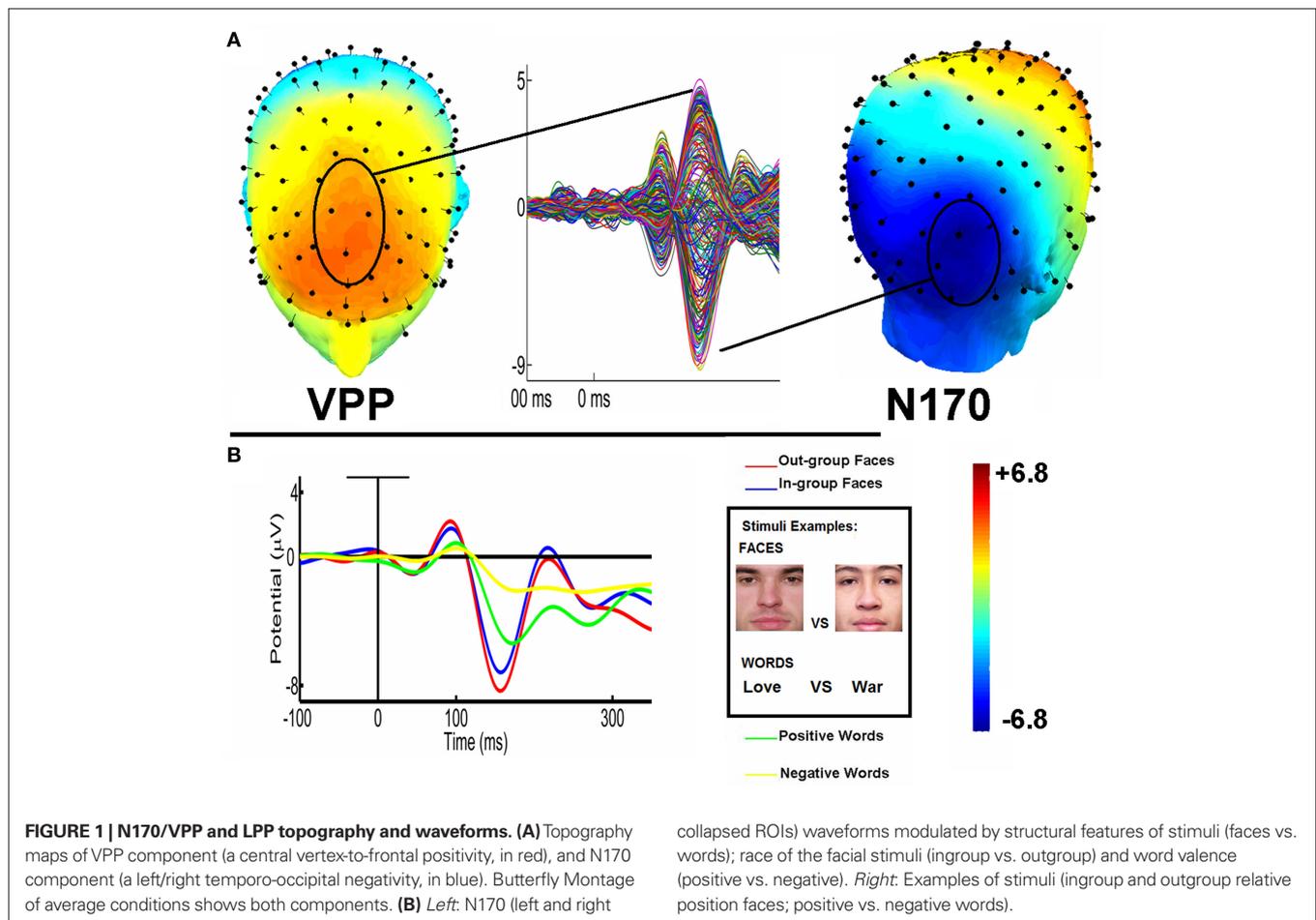
In the analysis of faces, no effect of intergroup relative position (IP) or contextual association of words was observed; an interaction between those factors was neither found. Regarding words, a valence effect revealed that positive words produced significantly higher amplitudes ($M = 3.13 \mu\text{V}$, $SD = 0.41$) than negative ones ($M = 2.62 \mu\text{V}$, $SD = 0.38$) [$F(1, 32) = 7.30$, $p = 0.01$]. In brief, the VPP component was higher for positive words than negative ones.

In order to compare the results not in terms of the relative ingroup/outgroup, but rather in terms of participants' group and race of the stimuli (indigenous or non-indigenous), the data were re-analyzed with the same number of factors (except for the fact that the race stimuli association was divided into indigenous vs. non-indigenous). Similar effects were observed; however, race was not significant for face [$F(1, 32) = 0.19$, $p = 0.67$], nor for words (context association, [$F(1, 32) = 0.23$, $p = 0.52$]), which indicates that race stimulus (indigenous or non-indigenous) on its own did not produce VPP modulation.

N170

For face stimuli, outgroup stimuli generated higher negative amplitudes ($M = -4.96 \mu\text{V}$, $SD = 0.49$) than ingroup ($M = -4.35 \mu\text{V}$, $SD = 0.47$). More importantly, we found an interaction of IP and ROI, [$F(1, 32) = 14.2$, $p < 0.001$]. *Post hoc* comparisons (Tukey HSD Test, $MS = 10.9$, $df = 36.7$) showed that the amplitudes for faces were more negative in the right region for outgroup stimuli (outgroup: $M = -6.32 \mu\text{V}$, $SD = 0.41$, ingroup: $M = -4.36 \mu\text{V}$, $SD = 0.38$, $p < 0.01$). In addition, an interaction of IP, contextual association, and ROI were all significant as well [$F(1, 32) = 9.32$, $p < 0.01$]. *Post hoc* comparisons showed that when associated with positive-valence words, outgroup faces produced more negative amplitude ($M = -6.46 \mu\text{V}$, $SD = 0.50$) than ingroup faces ($M = -5.61 \mu\text{V}$, $SD = 0.63$; $p < 0.001$).

For word stimuli, positive words generated more negative amplitudes ($M = -4.30 \mu\text{V}$, $SD = 0.67$) than negative words (valence effect, $M = -3.11 \mu\text{V}$, $SD = 0.47$). There was also an interaction between valence, context association, and ROI. *Post hoc*



comparisons (Tukey HSD test, $MS = 23.0$, $df = 40.4$) revealed that in the left region, positive words associated with ingroup stimuli, ($M = -3.37 \mu\text{V}$, $SD = 0.39$) differed from negative-valence words ($M = -1.93 \mu\text{V}$, $SD = 0.34$, $p < 0.001$). There were no semantic valence effects when words were associated to outgroup stimuli ($p = 0.15$).

In order to determine how the results behave if factor IP of the stimuli was computed on the basis of absolute social categories of faces (i.e., indigenous vs. non-indigenous instead of the relative categories of ingroup vs. outgroup) the mixed ANOVA was implemented again using the same factors for faces and words. Although several factors showed significant effects, the main difference between this ANOVA and the previous one was that the stimuli race association variable did not show an effect on faces [$F(1, 32) = 0.21$, $p = 0.65$] nor words [$F(1, 32) = 0.16$, $p = 0.87$]. Therefore, the race of stimuli association *per se* was not significant.

In summary, amplitudes of N170 component for faces were higher in response to outgroup than ingroup faces. In the right hemisphere, outgroup vs. ingroup faces were more differentiated when associated with positive words. For words, positive-valence elicited more negative amplitude than negative-valence words, and those differences were observed specially in the contextual association to outgroup faces. Words showed increased modulation in the

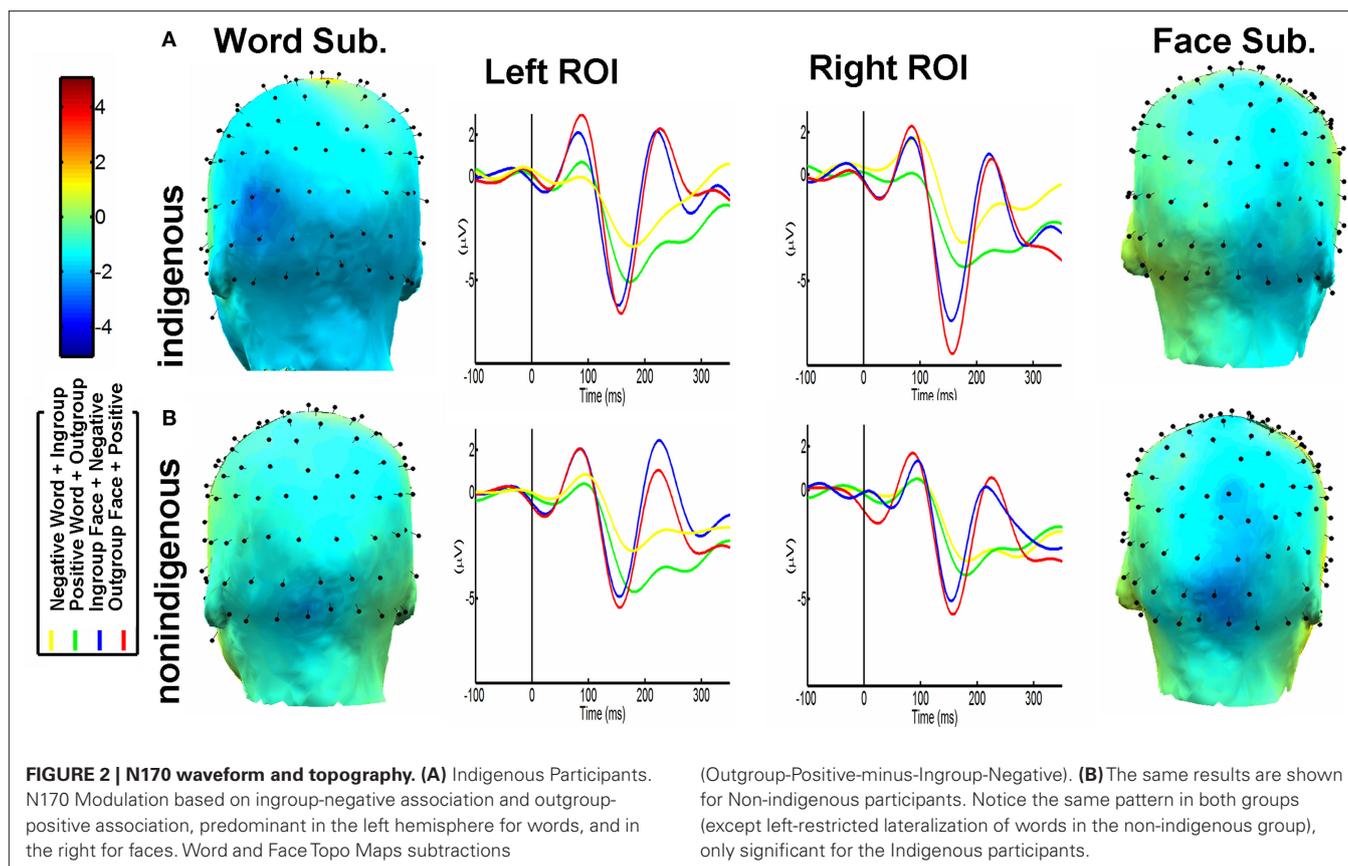
left region. In this hemisphere, positive-valence words were clearly different from negative-valence words when they were associated with ingroup stimuli.

In order to analyze the specific effects for indigenous and non-indigenous participants, the N170 component from each group was analyzed separately.

INDIGENOUS PARTICIPANTS

Faces

The ANOVA conducted revealed significant main effects of IP [outgroup > ingroup, $F(1, 16) = 4.83$, $p = 0.043$], and ROI [left < right, $F(1, 16) = 8.15$, $p = 0.011$]. Likewise, interaction effects of IP, ROI and contextual association [$F(1, 16) = 8.67$, $p < 0.05$] were observed. **Figure 2A** shows this last interaction effect. Outgroup faces associated with positive words and ingroup faces associated with negative words showed maximum difference in the right ROI. *Post hoc* comparisons (Tukey HSD Test, $MS = 9.65$, $df = 16.00$), confirmed the differences observed: indigenous faces elicited less negative amplitude in the right region associated with negative valence ($M = -3.88 \mu\text{V}$, $SD = 0.62$) compared to non-indigenous faces associated with positive valence ($M = -6.11 \mu\text{V}$, $SD = 0.54$; $p < 0.001$). Furthermore, there were no significant differences when the contextual association pattern (ingroup-positive vs. outgroup-negative) was reversed.



Words

Effects of Valence [positive > negative, $F(1, 16) = 25.4, p < 0.0001$] and ROI [left > right, $F(1, 16) = 7.34, p < 0.5$] were observed. Like face comparisons, an interaction between valence, ROI and contextual association was observed [$F(1, 16) = 4.67, p = 0.04$, see **Figure 2A**]. *Post hoc* comparisons over this interaction (Tukey HSD Test) evidenced a bilateral effect (but more prominent for the left region): positive words associated with outgroup ($M = -3.98 \mu\text{V}$, $SD = 0.45$) elicited significant differences with negative words associated with ingroup contexts ($M = -2.28 \mu\text{V}$, $SD = 0.23$; $p = 0.01$). No contextual association differences were observed regarding the opposite pattern (positive-ingroup vs. negative-outgroup).

NON-INDIGENOUS PARTICIPANTS

Faces

Contrary to indigenous participants, no significant differences were obtained for ROI [left < right, $F(1, 16) = 0.56, p = 0.47$], nor IP [outgroup > ingroup association, $F(1, 16) = 1.62, p = 0.22$]. Although non-indigenous faces linked with negative valence showed less negative amplitude ($M = -4.34 \mu\text{V}$, $SD = 0.67$) than indigenous faces associated with positive valence ($M = -4.85 \mu\text{V}$, $SD = 0.59, p = 0.068$) in the right ROI, no interaction between IP, contextual association, and ROI was observed [$F(1, 16) = 1.50, p = 1.02$, see **Figure 2B**].

Words

The ANOVA showed a significant main effect of valence [$F(1, 16) = 15.0, p = 0.001$, positive > negative]. In addition, similar to indigenous participants, a trend to significance was shown for

the interaction between valence, contextual association, and ROI [$F(1, 16) = 3.50, p = 0.06$]. **Figure 2B** shows said latter trend. As in the case of indigenous participants (reversed by ingroup/outgroup effect), *post hoc* comparisons performed on this interaction (Tukey HSD Test, $MS = 5.12, df = 16.0$) revealed lower amplitudes in the left ROI for negative words associated with non-indigenous targets ($M = -1.98 \mu\text{V}$, $SD = 0.29$) compared to positive words associated with indigenous targets ($M = -3.78 \mu\text{V}$, $SD = 0.31, p < 0.01$). It is important to note, as in the case of indigenous participants, that by reversing the contextual association pattern (positive/ingroup and negative/outgroup), no significant differences were observed.

DISCUSSION

BEHAVIORAL DATA

Among the IAT scores, the significantly higher reaction times corresponded to the indigenous group in the compatible task biased toward the ingroup. When the responses to images and words were analyzed separately, word scores were different from zero, significant for the indigenous group and close to significant for the non-indigenous group. This result is consistent with previous research on IAT, in which words presented more accentuated IAT scores than faces (Foroni and Bel-Bahar, 2009). In other words, the indigenous group showed strong ingroup favoritism compared with non-indigenous participants. These results are not inconsistent with previous reports of contextual malleability of implicit attitudes (Dasgupta and Greenwald, 2001; Rudman et al., 2001; Blair, 2002). In fact and more important, our results are in line with previous reports of IAT measures in Chilean populations (Hurtado et al., 2009).

Minorities perceived as belonging to a lower status (as in the case of our indigenous sample) are expected to have strong ingroup bias (Mullen et al., 1992; Reicher, 2004). Ingroup bias seems to increase with group salience and slightly decrease with social status (Mullen et al., 1992). Ingroup bias should be stronger in the indigenous group compared to the non-indigenous group, which is positioned as a higher-status majority.

ERPs RESULTS

Interestingly, the VPP was modulated by the stimulus' valence association, revealing that words with positive valence generated more positive amplitudes than negative-valence words. This result suggests that valence is processed before complete and late semantic processing (e.g., N400, semantic LPP components). This report, along with others which have found similar early semantic processing (Dien et al., 2003; Proverbio et al., 2007; Segalowitz and Zheng, 2009), is consistent with a timeline of word recognition in which semantic analysis, including the evaluation of emotional quality, occurs at an early stage of processing (Scott et al., 2009). Furthermore, our study shows an effect of functional dissociation between the VPP and N170 components: VPP was not modulated by the blending effects between race of the stimulus and valence of the word, which did occur – lateralized – for words and faces in N170. Although functional similarities between VPP and N170 have been proposed (Rossion et al., 2000; Bentin et al., 1996; Jeffreys, 1996; Itier and Taylor, 2002); in this study (see also Rossion et al., 2003; Joyce and Rossion, 2005), the complex effects observed in the N170 disappear in the VPP (probably due to the absence of VPP lateralization).

Regarding N170, Our results replicated previously reported (1) lateralization effects (increased amplitude for faces in the right, and for words in the left (Maurer et al., 2005); (2) faces' social category (outgroup faces modulate amplitude of the early component (Ito and Urland, 2003; Herrmann et al., 2007; Stahl et al., 2008; but see Ito and Urland, 2005 for opposite results); (3) N170 modulation by emotional effects, showing increased amplitude from positive stimuli over neutral stimuli (Schacht and Sommer, 2009); and finally, (4) modulation by contextual clues (Galli et al., 2006). In the left region, positive-valence words associated with ingroup stimuli were clearly distinguished from negative valence. These differences were not observed for outgroup stimuli. In the same region, there were no differences between face stimuli. As in the VPP results, these effects suggest that words can be processed as a whole in a global visual pattern, within a more holistic process (e.g., Simon et al., 2007), probably accelerated by emotional content. In the right region, faces generated more negative amplitude for outgroup stimuli, especially in the contextual association with positive words. Words did not show significant differences between subcategories in the right region. Words and faces also had a lateralized effect in outgroup-positive and ingroup-negative contextual associations: (1) Statistical differences between negative words associated with ingroup context vs. positive words associated with outgroup context to the left; and (2) statistical amplitude differences in the right region for ingroup faces associated with negative valence vs. outgroup faces associated with positive valence. Nevertheless, non-indigenous participants did not exhibit a main effect of social category association, and showed only a trend to

significance for the interaction valence association, IP, and ROIs. In both participant groups, when the contextual association pattern was reversed (positive and ingroup; negative and outgroup in both faces and words), the differences were canceled out. Therefore, N170 effects not only revealed modulation of valence and race association at early structural processing, but they also suggest the contextual integration of both effects (valence of words and relative social position of faces).

In a recent study, we reported the late ERPs (LPP) obtained in response to the same paradigm (Hurtado et al., 2009). In accordance with behavioral measures drawn from the IAT, we had shown a frontal LPP discrimination of semantic content based on compatible blocks with prejudice against indigenous targets in the word stimulus condition only in indigenous participants, which supports previously found lateralization results for positive–negative words. This LPP was indicative of the frontal distribution that had been previously linked to evaluative categorization. No effects of contextual blending of valence and membership were found in LPP. In the present report, we found an early effect of contextual blending, which is suggestive of an early integration of valence and race of stimuli. This early effect of contextual blending of valence and membership as well as the late effect of arousal and higher order cognitive processes are in accordance with comparable results from other paradigms. For example, valence and several other features of the stimuli seem to be processed early, but the higher order composite significance seems to be processed rather later (Codispoti et al., 2007; Olofsson et al., 2008; Rozenkrants and Polich, 2008). Along with this background, taken together, both studies of ERP/IAT indicate that the configuration aspects of stimuli (race and valence) are discriminated early, but their combined effects (compatible or incompatible associations) are processed at later phases.

OVERALL RESULTS

In addition to previous reports of left/right lateralization for words and faces in N170, we found an effect of early contextual blending of valence (words) and social membership features (faces). To the extent of our knowledge, this is the first study that presents IAT early brain correlates, suggesting that valence and race processing may occur not only before or parallel to the face structural coding process (Ashley et al., 2004), but also in a discriminated and combined manner. In both groups, ERPs are consistent in evidencing a predominance of ingroup distinction against outgroup. N170 discriminates word valence only at the ingroup level, and negative ingroup categories against positive outgroup categories produce the major difference. These convergent results suggest that semantic valence associations have a particular influence when the distinction between ingroup and outgroup is at stake. As for the differences between indigenous and non-indigenous participants, a convergence between behavioral and ERPs measures was also observed.

In general, indigenous participants exhibited more ingroup bias, and therefore various measures showed a differential pattern in that direction. On the IAT scores, only indigenous participants showed significant ingroup bias in general and word measures. In this group, N170 showed the main effect of the outgroup-ingroup distinction, and significant differences for positive/ingroup and negative/outgroup associations for word (left ROI) and faces (right ROI); yet, this was not significant among the non-indigenous

participants. Such bias toward the ingroup is consistent with fMRI studies which suggest emotion-biased motivated reasoning (Hamann, 2003; Westen et al., 2006).

The present research work has two potential confounds. First, indigenous participants were tested on the university campus. Considering that it may be a place they would perceive as an “out-group place” in and of itself, it is likely that, when testing indigenous participants, the social context plays a different role than for the non-indigenous participants, who were students and professors. Those social context factors may perhaps explain why indigenous participants presented a more accentuated IAT effect and bigger N170 modulation. Since Chilean indigenous members do not live segregated, as well as due to infrastructure limitations (the lab is situated on the university campus), it was not possible to test this potential social context effect on IAT measures and ERPs. Second, we used 16 blocks instead of 8 blocks, as in the original IAT. Although this represents a modification of the classic IAT paradigm, it was designed in order to control for the influence of the dominant hand responses, reason why we counterbalanced the hand assignments in order to obtain all possible combinations of stimulus types and responses.

IMPLICATIONS FOR ATTITUDES RESEARCH

This is the first report of early electrophysiological correlates of IAT, which is novel and relevant in the context of social psychology. First, significant relationships were observed between ERPs and IAT consistent with the expected patterns of group bias. In other words, there was a systematic modulation of electrophysiological processes and implicit behavioral responses in the IAT task. This suggests that, early on, there is a reaction to stimulation that is semantically unfavorable to the group itself, and said reaction is coherent with the difficulty shown by the subjects in tasks that involve working with such stimuli. The results indicate that a crucial aspect of the IAT cognitive task lies in processing (relationships of) stimuli that are contrary to a positive representation of one's own social group. Furthermore, it is possible that recognition or any form of processing of stimulation semantically unfavorable to the target group represents an essential clue to understand the dynamics of prejudice, as well as the expression of attitudes in intergroup contexts in general. This would be related to pessimistic bias research (Kaffenberger et al., 2010), demonstrating changed activations in brain areas known to be involved in processing negative emotions.

The results of our study are in accordance with recent ERPs reports of ingroup/outgroup research, revealing an inverted pattern of amplitude based on participants' relative social position regarding ingroup/outgroup stimuli. Dickter and Bartholow (2007) used an adapted version of the Eriksen flanker task, showed an ERP reverse pattern between groups of participants (ingroup vs. outgroup participants). In our recent study using this paradigm (Hurtado et al., 2009), we reported an correspondence between the degree of ingroup bias (IAT results) and lateralization of LPP amplitude: a more racial bias effect and a stronger LPP lateralization based on compatible/incompatible blocks. The study showed the same reversal pattern between groups based on membership. Our present work corroborates the role of membership position in the reversion of ERP patterns. Altogether, the findings confirm that

physical features of faces and the semantic valence association *per se* do not affect the results, but that their effect is manifested in the dependence of the relation between the stimulus and the perceiver's ingroup membership.

The fast growing field of social neuroscience (Adolphs, 2003; Frith and Frith, 2010) promotes an integration of the methodologies and approaches of social psychology and neuroscience. Our results, using combined methodology (IAT and ERPs), suggest a strong racial bias effect at early neural processing, showing a convergence between behavioral and electrophysiological measures. However, this effect of prejudice over neural responses should be taken with caution. The IAT has been questioned from several perspectives. Some authors have argued that cognitive-emotional associations should not be conceptualized as implicit prejudice (Arkes and Tetlock, 2004). Those authors suggest that associations between a specific racial group and a category can be more easily associated, without it implying a negative effect. Kinoshita and Peek-O'Leary (2005, 2006) consider that a racial IAT effect can be interpreted as a “salience asymmetry account,” in which a participant is focused more on the familiar category. Hence simplifying the task (just like a simple detection task), pointing out that the IAT may (partially) reflect differential familiarity with the groups. In addition, individual differences in emotional processing, well known in neuroscience studies (Hamann and Canli, 2004), should have an effect on the IAT responses (see too Blanton et al., 2009). Even so, it has been argued that implicit measures of social cognition do not reflect only automatic processes but rather the joint contributions of multiple, qualitatively different processes (Conrey et al., 2005). In the same line, although implicit association tasks are robust measurements of implicit (unconscious) evaluation, the implicitness property has to be carefully considered. We adopt the De Houwer definition of implicit measure as “outcome of a measurement procedure that results from automatic processes by which the to-be measured attribute causally determines the outcome” (De Houwer et al., 2009). Our task can only partially assess the normative criterion of this definition. Given that the N170 does not directly depend on explicit activity or awareness, we can consider the N170 modulation as an implicit measurement. Future studies, with an experimental design able to compare the modulation of the N170 contextual effects by implicit (unconsciousness) and explicit (consciousness) tasks will allow establishing the differential effects between both kinds of measurements. Probably, top-down emotional control will attenuate emotional arousal (see Herwig et al., 2010).

Our findings reveal that there are several processes interacting at the early stages in various ways, and probably under a dynamic pattern of association, which are present for different combinations of stimuli. The contribution of electrophysiological and multi-leveled studies allows us to advance toward a dynamic conception of racial bias. In this conception, many basic dynamically-formed subprocesses such as the interaction between stimulus familiarity and kinship, individual strategy tasks, and different semantic and structural process combinations take part to feed complex processes like prejudice (Ibáñez et al., 2009). Assessing such possible differential effects using ERPs and IAT calls for further inquiries.

In general, these lines of implications of our results are consistent with recent reports of cognitive context-dependence effects (Ibáñez et al., 2006, 2009, 2010a,b,c; Cornejo et al., 2009; Hurtado et al., 2009; Aravena et al., 2010; Riveros et al., 2010; San Martín et al., 2010), and more precisely, with research on the contextual malleability of implicit attitudes (Dasgupta and Greenwald, 2001; Rudman et al., 2001; Blair, 2002). For decades, attitudes toward groups were assumed to be stable and context-independent knowledge structures. However, in recent years, a growing body of evidence has suggested that these knowledge structures are considerably more flexible than originally assumed. In particular, previous research on the malleability of the IAT has shown that the context in which IAT scores are obtained affects these scores to a great extent (Mierke and Klauer, 2003; Mitchell et al., 2003). The implications of this idea is that implicit attitudes should not be conceptualized and measured as if they were stable associations stored in long term memory, but rather, on-the-fly constructions of behavior. According to this conceptualization, attitudes may be understood as temporary constructs that change with the context in which they are made. If so, then it should be possible to reverse the relative impact of a person's minority and majority identities on judgment by subtle variations of the judgment context.

REFERENCES

- Adolphs, R. (2003). Investigating the cognitive neuroscience of social behavior. *Neuropsychologia* 41, 119–126.
- Allport, G. W. (1954). *The Nature of Prejudice*. Reading, MA: Addison Wesley.
- Aravena, P., Hurtado, E., Riveros, R., Cardona, F., Manes, F., and Ibáñez, A. (2010). Applauding with closed hands: neural signature of action sentence compatibility effects. *PLoS ONE* 5, e11751. doi: 10.1371/journal.pone.0011751.
- Arkes, H. R., and Tetlock, P. E. (2004). Attributions of implicit prejudice, or “would Jesse Jackson ‘fail’ the implicit Association Test?” *Psychol. Inq.* 15, 257–278.
- Ashley, V., Vuilleumier, P., and Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *Neuroreport* 15, 211.
- Bargh, J. A., (1997). The automaticity of everyday life. *Adv. Soc. Cogn.* 10, 1–61.
- Batty, M., and Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cogn. Brain Res.* 17, 613–620.
- Bengoa, J. (2000). *Historia del Pueblo Mapuche: (siglo XIX y XX)*. Santiago: Lom Ediciones.
- Bentin, S., Allison, T., Puce, A., Perez, E., and McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Pers. Soc. Psychol. Rev.* 6, 242.
- Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., and Tetlock, P. E. (2009). Strong claims and weak evidence: reassessing the predictive validity of the IAT. *J. Appl. Psychol.* 94, 567–582; discussion 583–603.
- Blau, V. C., Maurer, U., Tottenham, N., and McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behav. Brain Funct.* 3, 7.
- Bohner, G., Siebler, F., González, R., Haye, A., and Schmidt, E. (2008). Situational flexibility of ingroup-related attitudes: a single category IAT study of people with dual national identity. *Group Process. Intergroup Relat.* 11, 301–317.
- Brandeis, D., Lehmann, D., Michel, C. M., and Mingrone, W. (1995). Mapping event-related brain potential microstates to sentence endings. *Brain Topogr.* 8, 145–159.
- Brown, R., González, R., Zagefka, H., Manzi, J., and Cehajic, S. (2008). Nuestra culpa: collective guilt and shame as predictors of reparation for historical wrongdoing. *J. Pers. Soc. Psychol.* 94, 75.
- Caldara, R., Rossion, B., Bovet, P., and Hauert, C. A. (2004). Event-related potentials and time course of the ‘other-race’ face classification advantage. *Neuroreport* 15, 905.
- Caldara, R., Thut, G., Servois, P., Michel, C. M., Bovet, P., and Renault, B. (2003). Face versus non-face object perception and the ‘other-race’ effect: a spatio-temporal event-related potential study. *Clin. Neurophysiol.* 114, 515–528.
- Codispoti, M., Ferrari, V., and Bradley, M. M. (2007). Repetition and event-related potentials: distinguishing early and late processes in affective picture perception. *J. Cogn. Neurosci.* 19, 577–586.
- Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., and Groom, C. J. (2005). Separating multiple processes in implicit social cognition: the quad model of implicit task performance. *J. Pers. Soc. Psychol.* 89, 469–487.
- Cornejo, C., Simonetti, F., Ibáñez, A., Aldunate, N., Lopez, V., Ceric, F., and Nuñez, R. (2009). Gesture and metaphor: electrophysiological evidence of 400 multimodal modulation. *Brain Cogn.* 70, 42–52.
- Dasgupta, N., and Greenwald, A. G. (2001). On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *J. Pers. Soc. Psychol.* 81, 800–814.
- Dasgupta, N., McGhee, D. E., Greenwald, A. G., and Banaji, M. R. (2000). Automatic preference for White Americans: eliminating the familiarity explanation. *J. Exp. Soc. Psychol.* 36, 316–328.
- De Houwer, J., Teige-Mocigemba, S., Spruyt, A., and Moors, A. (2009). Implicit measures: a normative analysis and review. *Psychol. Bull.* 135, 347–368.
- Devine, P. G. (1989). Stereotypes and prejudice: their automatic and controlled components. *J. Pers. Soc. Psychol.* 56, 5–18.
- Dickter, C. L., and Bartholow, B. D. (2007). Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Soc. Cogn. Affect. Neurosci.* 2, 189–198.
- Dien, J., Frishkoff, G. A., Cerbone, A., and Tucker, D. M. (2003). Parametric analysis of event-related potentials in semantic comprehension: evidence for parallel brain mechanisms. *Cogn. Brain Res.* 15, 137–153.
- Dillehay, T. D. (2007). *Monuments, Empires, and Resistance: The Araucanian Polity and Ritual Narratives*. Cambridge: Cambridge University Press.
- Dunham, Y., Baron, A. S., and Banaji, M. R. (2006). From American city to Japanese village: a cross-cultural investigation of implicit race attitudes. *Child Dev.* 77, 1268.
- Efferson, C., Lalive, R., and Fehr, E. (2008). The coevolution of cultural groups and ingroup favoritism. *Science* 321, 1844.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: an overview. *Cogn. Emot.* 15, 115–141.
- Fazio, R. H., and Olson, M. A. (2003). Implicit measures in social cognition research: their meaning and use. *Annu. Rev. Psychol.* 54, 297–327.
- Foroni, F., and Bel-Bahar, T. (2009). Picture-IAT versus Word-IAT: level of stimulus representation influences on the IAT. *Eur. J. Soc. Psychol.* 40, 321–337.
- Frith, C., and Frith, U. (2010). Learning from others: introduction to the

CONCLUSION

Our results suggest that implicit association effects influence early stages of processing. A neural discrimination of combined association between word valence and face's relative social position was observed, showing a different modulation pattern depending on the combination of both variables (relative social position and valence association). Accordingly, when behavioral measures of ingroup favoritism were stronger (indigenous participants), an increased early contextual modulation of ERPs was observed, suggesting an early influence of contextual effects on IAT measures. In other words, these results suggest that high level contextual combinations of social relative position and valence have an early neural modulation and a consequent behavioral effect. The assessment of IAT simultaneously with ERP recordings opens a new path for the interplay of social attitude research and social neuroscience.

ACKNOWLEDGMENTS

This research was partially supported by a CONICET grant and by a UDP Neuroscience Grant CG090325006. We wish to thank Alvaro Navarro for his technical support. We are indebted to Dr. David Amodio for his helpful comments in a previous version of this manuscript.

- special review series on social neuroscience. *Neuron* 65, 739–743.
- Galli, G., Feurra, M., and Viggiano, M. P. (2006). “Did you see him in the newspaper?” Electrophysiological correlates of context and valence in face processing. *Brain Res.* 1119, 190–202.
- Gawronski, B., Hofmann, W., and Wilbur, C. J. (2006). Are “implicit” attitudes unconscious? *Conscious. Cogn.* 15, 485–499.
- Gawronski, B., Peters, K. R., Brochu, P. M., and Strack, F. (2008). Understanding the relations between different forms of racial prejudice: a cognitive consistency perspective. *Pers. Soc. Psychol. Bull.* 34, 648.
- Glick, P., and Fiske, S. T. (2001). An ambivalent alliance: hostile and benevolent sexism as complementary justifications for gender inequality. *Am. Psychol.* 56, 109–118.
- Gobbini, M. I., and Haxby, J. V. (2006). Neural response to the visual familiarity of faces. *Brain Res. Bull.* 71, 76–82.
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *J. Pers. Soc. Psychol.* 74, 1464–1480.
- Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the implicit Association Test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85, 197–216.
- Hamann, S. (2003). Nosing in on the emotional brain. *Nat. Neurosci.* 6, 106–108.
- Hamann, S., and Canli, T. (2004). Individual differences in emotion processing. *Curr. Opin. Neurobiol.* 14, 233–238.
- Haye, A., González, R., Ordóñez, G., Bohner, G., Siebler, F., Sirlopú, D., Millar, A., De Tezanos-Pinto, P., and Torres, D. (2010). System-perpetuating asymmetries between explicitly implicit intergroup attitudes among indigenous and non-indigenous Chileans. *Asian J. Soc. Psychol.* 13, 163–172.
- He, Y., Johnson, M. K., Dovidio, J. F., and McCarthy, G. (2009). The relation between race-related implicit associations and scalp-recorded neural activity evoked by faces from different races. *Soc. Neurosci.* 4, 426–442.
- Herrmann, M. J., Schreppe, T., Jäger, D., Koehler, S., Ehlis, A. C., and Fallgatter, A. J. (2007). The other-race effect for face perception: an event-related potential study. *J. Neural Transm.* 114, 951–957.
- Herwig, U., Kaffenberger, T., Jäncke, L., and Brühl, A. B. (2010). Self-related awareness and emotion regulation. *Neuroimage* 50, 734–741.
- Hillyard, S. A., and Picton, T. W. (1987). “Electrophysiology of cognition,” in *Handbook of Physiology, Section 1: The Nervous System. Volume V. Higher Functions of the Brain, Part 2*, ed F. Plum (Bethesda, MD: American Physiological Society), 519–584.
- Honda, Y., Watanabe, S., Nakamura, M., Miki, K., and Kakigi, R. (2007). Interhemispheric difference for upright and inverted face perception in humans: an event-related potential study. *Brain Topogr.* 20, 31–39.
- Horowitz, S. G., Rossion, B., Skudlarski, P., and Gore, J. C. (2004). Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. *Neuroimage* 22, 1587–1595.
- Hurtado, E., Haye, A., González, R., Manes, F., and Ibáñez, A. (2009). Contextual blending of ingroup/outgroup face stimuli and word valence: LPP modulation and convergence of measures. *BMC Neurosci.* 10, 69. doi: 10.1186/1471-2202-10-69.
- Ibáñez, A., Haye, A., González, R., Hurtado, E., and Henríquez, R. (2009). Multi-level analysis of cultural phenomena: the role of ERP approach to prejudice. *J. Theory Soc. Behav.* 39, 81–110.
- Ibáñez, A., Riveros, R., Aravena, P., Vergara, V., Cardona, J. F., García, L., Hurtado, E., and Martin Reyes, M., Barutta, J., and Manes, F. (2010a). When context is hard to integrate: cortical measures of congruency in schizophrenics and healthy relatives from multiplex families. *Schizophr. Res.* doi: 10.1016/j.schres.2010.04.008. http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6TC2-50338T7-1&_user=10&_coverDate=05%2F15%2F2010&_rdoc=1&_fmt=high&_orig=search&_origin=search&_sort=d&_docanchor=&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=be0f8726a61c22d5ac2377f59de8f31&searchtype=a
- Ibáñez, A., Manes, F., Escobar, J., Trujillo, N., Andreucci, P., and Hurtado, E. (2010b). Gesture influences the processing of figurative language in non-native speakers. *Neurosci. Lett.* 471, 48–52.
- Ibáñez, A., Toro, P., Cornejo, C., Urquina, H., Manes, F., Weisbrod, M., and Schröder, J. (2010c). High contextual sensitivity of metaphorical expressions and gesture blending: a video ERP design. *Psychiatry Res. Neuroimaging*. doi: 10.1016/j.psychres.2010.08.008. (In press).
- Ishizu, T., Ayabe, T., and Kojima, S. (2008). Configurational factors in the perception of faces and non-facial objects: an ERP study. *Int. J. Neurosci.* 118, 955–966.
- Itier, R. J., and Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *Neuroimage* 15, 353–372.
- Ito, T. A., and Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *J. Pers. Soc. Psychol.* 85, 616–626.
- Ito, T. A., and Urland, G. R. (2005). The influence of processing objectives on the perception of faces: an ERP study of race and gender perception. *Cogn. Affect. Behav. Neurosci.* 5, 21–36.
- Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Vis. Cogn.* 3, 1–38.
- Jemel, B., Coutya, J., Langer, C., and Roy, S. (2009). From upright to upside-down presentation: a spatio-temporal ERP study of the parametric effect of rotation on face and house processing. *BMC Neurosci.* 10, 100. doi: 10.1186/1471-2202-10-100.
- Joyce, C., and Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clin. Neurophysiol.* 116, 2613–2631.
- Kaffenberger, T., Brühl, A. B., Baumgartner, T., Jäncke, L., and Herwig, U. (2010). Negative bias of processing ambiguously cued emotional stimuli. *Neuroreport* 21, 601–605.
- Kinoshita, S., and Peek-O’Leary, M. (2005). Does the compatibility effect in the race implicit Association Test reflect familiarity or affect? *Psychon. Bull. Rev.* 12, 442.
- Kinoshita, S., and Peek-O’Leary, M. (2006). Two bases of the compatibility effect in the Implicit Association Test (IAT). *Q. J. Exp. Psychol.* 59, 2102.
- Krombholz, A., Schaefer, F., and Boucsein, W. (2007). Modification of N170 by different emotional expression of schematic faces. *Biol. Psychol.* 76, 156–162.
- Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., and Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neurosci. Lett.* 253, 147–150.
- Marzi, T., and Viggiano, M. P. (2007). Interplay between familiarity and orientation in face processing: an ERP study. *Int. J. Psychophysiol.* 65, 182–192.
- Maurer, U., Brandeis, D., and McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behav. Brain Funct.* 9, 13.
- Mierke, J., and Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *J. Pers. Soc. Psychol.* 85, 1180–1192.
- Mitchell, J. P., Nosek, B. A., and Banaji, M. R. (2003). Contextual variations in implicit evaluation. *J. Exp. Psychol. Gen.* 132, 455–469.
- Mullen, B., Brown, R., and Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: an integration. *Eur. J. Soc. Psychol.* 22, 103–122.
- Nosek, B. A., and Banaji, M. R. (2001). The go/no-go association task. *Soc. Cogn.* 19, 625–666.
- Oken, B. S., and Chiappa, K. H. (1986). Statistical issues concerning computerized analysis of brain wave topography. *Ann. Neurol.* 19, 493–494.
- Olofsson, J. K., Nordin, S., Sequeira, H., and Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biol. Psychol.* 77, 247–265.
- Pizzagalli, D. A., Lehmann, D., Hendrick, A. M., REGARD, M., Pascual-Marqui, R. D., and Davidson, R. J. (2002). Affective judgments of faces modulate early activity (160 ms) within the fusiform gyri. *Neuroimage* 16, 663–677.
- Pourtois, G., Grandjean, D., Sander, D., and Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cereb. Cortex* 14, 619–633.
- Proverbio, A. M., Del Zotto, M., and Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indices of animal vs. artifact recognition. *BMC Neurosci.* 8, 24. doi: 10.1186/1471-2202-8-24.
- Rebai, M., Poiroux, S., Bernard, C., and Lalonde, R. (2001). Event-related potentials for category-specific information during passive viewing of faces and objects. *Int. J. Neurosci.* 106, 209–226.
- Reicher, S. (2004). The context of social identity: domination, resistance, and change. *Polit. Psychol.* 25, 921–945.
- Riveros, R., Manes, F., Hurtado, E., Escobar, M., Reyes, M., Cetkovich, M., and Ibáñez, A. (2010). Context-sensitive social cognition is impaired in schizophrenic patients and their healthy relatives. *Schizophr. Res.* 116, 297–298.
- Rossion, B., Delvenne, J. F., Debatiste, D., Goffaux, V., Bruyer, R., Crommelinck, M., and Guérit, J. M. (1999). Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biol. Psychol.* 50, 173–189.
- Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., and Crommelinck, M. (2002). Expertise training with novel objects leads to left-lateralized face-like

- electrophysiological responses. *Psychol. Sci.* 13, 250–257.
- Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., and Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport* 11, 69.
- Rossion, B., and Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage* 39, 1959–1979.
- Rossion, B., Joyce, C. A., Cottrell, G. W., and Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage* 20, 1609–1624.
- Rothermund, K., and Wentura, D. (2004). Underlying processes in the implicit Association Test: dissociating salience from associations. *J. Exp. Psychol. Gen.* 133, 139–165.
- Rozenkrants, B., and Polich, J. (2008). Affective ERP processing in a visual odd-ball task: arousal, valence, and gender. *Clin. Neurophysiol.* 119, 2260–2265.
- Rudman, L. A., Ashmore, R. D., and Gary, M. L. (2001). “Unlearning” automatic biases: the malleability of implicit prejudice and stereotypes. *J. Pers. Soc. Psychol.* 81, 856–868.
- Sagiv, N., and Bentin, S. (2001). Structural encoding of human and schematic faces: holistic and part-based processes. *J. Cogn. Neurosci.* 13, 937–951.
- Saiz, J. L., and Williams, J. E. (1991). “Estereotipos del indígena mapuche: una verificación empírica de proposiciones y hallazgos previos,” *Sobre culturas indígenas: lenguaje e identidad*, eds En Y. Kuramochi and P. de la Peña (Temuco, Chile: CONICYT-UCT-UFRO), 141–148.
- San Martín, R., Manes, F., Hurtado, E., Isla, P., and Ibáñez, A. (2010). Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. *Neuroimage* 51, 1194–1204.
- Schacht, A., and Sommer, W. (2009). Emotions in word and face processing: early and late cortical responses. *Brain Cogn.* 69, 538–550.
- Scott, G. G., O'Donnell, P. J., Leuthold, H., and Sereno, S. C. (2009). Early emotion word processing: evidence from event-related potentials. *Biol. Psychol.* 80, 95–104.
- Segalowitz, S. J., and Zheng, X. (2009). An ERP study of category priming: evidence of early lexical semantic access. *Biol. Psychol.* 80, 122–129.
- Sibley, C. G., and Liu, J. H. (2007). New Zealand = bicultural? Implicit and explicit associations between ethnicity and nationhood in the New Zealand context. *Eur. J. Soc. Psychol.* 37, 1222.
- Simon, G., Petit, L., Bernard, C., and Rebai, M. (2007). N170 ERPs could represent a logographic processing strategy in visual word recognition. *Behav. Brain Funct.* 3, 21.
- Stahl, J., Wiese, H., and Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: an event-related potential study. *Neuroreport* 19, 583–587.
- Westen, D., Blagov, P. S., Harenski, K., Kiltz, C., and Hamann, S. (2006). Neural bases of motivated reasoning: an fMRI study of emotional constraints on partisan political judgment in the 2004 US presidential election. *J. Cogn. Neurosci.* 18, 1947–1958.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 June 2010; accepted: 22 September 2010; published online: 19 October 2010.

Citation: Ibáñez A, Gleichgerrcht E, Hurtado E, González R, Hays A and Manes FF (2010) Early neural markers of implicit attitudes: N170 modulated by intergroup and evaluative contexts in IAT. *Front. Hum. Neurosci.* 4:188. doi: 10.3389/fnhum.2010.00188

Copyright © 2010 Ibáñez, Gleichgerrcht, Hurtado, González, Hays and Manes. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

SUPPLEMENTARY MATERIAL

IAT STIMULI VALIDATION

Validation of linguistic stimuli

In order to validate word content, a questionnaire was used to determine how pleasant or unpleasant a list of 150 words was, which included moderate frequency of use words selected by using the Lifcach software. Fifty psychology students participated, the average age was 19.62 (SD = 3.33), 67.3% were female. The participants had to answer using a Likert-scale where 1 represented “very pleasant” and 7 was “very unpleasant.” An ANOVA of repeated measures was used to contrast the categorizations for the list of pleasant and unpleasant words. Significant differences were obtained for the categorization of both lists [$F(1, 73) = 25.16$, $p < 0.001$]. From the list of pleasant words, only those that were rated between 1 and 3 points were chosen (72 pleasant words were chosen, 3 rejected). From the list of unpleasant words, we only chose those rated between 5 and 7 (71 unpleasant words were chosen, 4 rejected). A validation questionnaire was then used in order to rule out possible *a priori* associations between the chosen words and the indigenous or non-indigenous ethnic groups. This questionnaire was answered by 50 psychology students; the average age was 19.32 years, 20% male. Participants used a 7-point scale, rating the association of a word with an ethnic group, “1” being an extreme association with indigenous ethnic group and “7,” an extreme association with non-indigenous ethnic group. Words were evaluated as neutral with respect to the ethnic group with an average score of 4.21 (SD = 0.50). There was no significant and constant correlation or association found in our sample of participants between a specific certain word and one of the ethnic groups. Nevertheless, as a protective measure, those words that on average were ranked below the 5th percentile (3.40; tendency to associate with indigenous group) or over the 95th percentile (5.03; tendency to associate with non-indigenous group) were not included. Thus, from the 143 words proposed, 9 were eliminated because they could be associated with indigenous people and 8 because they could be associated with non-indigenous people. As a result of these two preliminary studies, 126 words were used in the final set for research purposes.

Validation of pictorial stimuli

The non-indigenous photographs were of students and professors, while the indigenous ones were taken at an indigenous community close to Santiago, Chile. All participants signed an informed consent accepting that their photographs would be used for experimental purposes. The photographs were standardized based on size, brightness and intensity. Only the eyes, nose and mouth regions were displayed, in order to eliminate any possible distractions. A total of 324 photographs were taken, 159 of non-indigenous individuals and 165 indigenous individuals. One hundred seven volunteers, all university students, (average age = 19.32 years, 65.7% female) answered a questionnaire to validate the photographs previously obtained. On a 7-point scale, the ethnic group of face was assessed, where “1” was completely indigenous, and “7” completely non-indigenous; additionally, they had to answer from 1 to 7 how pleasant (1) or unpleasant (7) they found the photograph. They were asked to rate the emotion of the face between 1 – Neutral, 2 – Happy, 3 – Afraid, 4 – Angry. The

averages obtained for the previous items tended toward the center of the scale, as follows: ethnic group of face: 4.44 (SD = 1.99); pleasant: 4.27 (SD = 1.59); and emotion: 2.34 (SD = 1.10). The images rated below the 5th percentile (very pleasant) or over the 95th percentile (very unpleasant) were eliminated. Of the faces showing neutral emotion, 87.04% were rated in the non-indigenous and 80.70% in the indigenous group. These analyses helped to select the images more commonly identified with each category (indigenous or non-indigenous ethnic group), as well as images that were not affected by extreme values of pleasant or unpleasant valence and that represent a neutral emotion rated with an absolute percentage. The images that were to be included in the final set for the IAT study and the electrophysiological study were chosen from this set of images.

EXTRACTION OF IAT MEASURES: ALGORITHM PROPOSED BY GREENWALD ET AL. (2003)

The following steps describe the application of the algorithm proposed by Greenwald et al. (2003) to the IAT test results relevant to this study.

1. All the data for the compatible (3 and 4) and incompatible blocks (7 and 8) was obtained from both key assignments.
2. Trails with a response time over 1000 ms were eliminated.
3. For each of the four blocks, the reaction-time average for trials with a correct response was obtained.
4. A *pooled standard deviation* was calculated for all block 3 and 7 trials, and then another for blocks 4 and 8.
5. For each trial with an incorrect response, the reaction time was replaced by the corresponding block average (computed in step 3) plus 600 ms. This corresponds to a penalty applied to wrong responses.
6. Based on the resulting reaction times, both correct and incorrect trials, the average reaction time was calculated for each of the four blocks.
7. The average difference was calculated between block 7 and 3, and another difference between block 8 and 4.
8. Each average was divided by the corresponding *pooled standard deviation* (computed in step 4).
9. The IAT score is the average of the two values obtained in the previous step.

The result of this procedure was a number with an estimated value close to zero for subjects who did not show racial bias in the test. Positive values correspond to the detection of bias in favor of the indigenous group, while negative values showed bias in favor of the non-indigenous group. By the way it is calculated, the result is similar to Cohen’s *d* coefficient; hence it can be interpreted as an effect size measure.

In order to enhance the measurement, and also to simultaneously record reaction times to faces and words, as in the procedure proposed by Greenwald et al. (2003), this study considered two additional measures, which were obtained separately taking into account face and word trials. Therefore, it was possible to explore, in further detail, the connection between the IAT test’s behavioral information and the descriptors resulting from the extraction of ERPs, which have different waveforms for words and

faces. Separate analyses were done based on stimulus type (face or word), experimental group (indigenous or non-indigenous) and task (compatible or incompatible with prejudice toward minority); resulting in eight categories. Finally, in order to obtain accuracy for both groups in this IAT test, correct responses were divided into total number of responses with an accuracy score between 0 and 1. Accuracy is calculated based on compatible and incompatible blocks and the separate responses to faces and words.

P1 EFFECTS

Certainly, in a preliminary analysis, we found early effects for stimulus type discrimination (face vs. word) on early P1/N1, but no effects of contextual association and valence were observed. Since the early stimulus type effects have been already reported and because we were mainly focused on N170 as a specific marker of facial processing, these stimulus type effects on P1/N1 results were not included in the current version of the manuscript.