doi:10.1093/scan/nsq065 SCAN (2011) 6, 468–476

Neural correlates of the implicit association test: evidence for semantic and emotional processing

John K. Williams[1] and Jason R. Themanson[2]
[1]Biola University, and [2]Illinois Wesleyan University

The implicit association test (IAT) has been widely used in social cognitive research over the past decade. Controversies have arisen over what cognitive processes are being tapped into using this task. While most models use behavioral (RT) results to support their claims, little research has examined neurocognitive correlates of these behavioral measures. The present study measured event-related brain potentials (ERPs) of participants while completing a gay-straight IAT in order to further understand the processes involved in a typical group bias IAT. Results indicated significantly smaller N400 amplitudes and significantly larger LPP amplitudes for compatible trials than for incompatible trials, suggesting that both the semantic and emotional congruence of stimuli paired together in an IAT trial contribute to the typical RT differences found, while no differences were present for earlier ERP components including the N2. These findings are discussed with respect to early and late processing in group bias IATs.

Keywords: N400; LPP; IAT; social cognition

INTRODUCTION

In the past decade, the implicit association test (IAT) has been widely used to measure underlying implicit attitudes related to a variety of constructs. While early studies were designed to measure prejudice toward different racial groups (Greenwald *et al.*, 1998) the IAT has since been used to measure implicit biases and preferences toward religious groups (Rowatt *et al.*, 2005), political candidates (Arcuri *et al.*, 2008), sexual orientation (Banse *et al.*, 2001; Boysen *et al.*, 2006) and gender roles (White and White, 2006). More recently, the IAT has also been used to measure personal attributes, such as implicit self-esteem (Greenwald and Farnham, 2000), humility (Rowatt *et al.*, 2006) and anxiety responses of phobics to spiders and snakes (Teachman *et al.*, 2001).

Along with the ubiquity of the IAT in social cognitive research have come a number of criticisms focused on its usefulness and validity for measuring implicit attitudes. While arguments go back and forth between critics and designers of the IAT as to its overall validity (Blanton and Jaccard, 2006; Greenwald *et al.*, 2006), others have sought to explain the sources of the response time differences found in typical IATs. One account argues that the IAT indicates the strength of semantic connections or associations between concepts (Greenwald *et al.*, 1998). An alternative proposal suggests a salience asymmetry explanation in which the familiarity of a concept makes it more or less salient, and concepts are connected together based on their relative salience to the participant (Rothermund and Wentura, 2004; Kinoshita and Peek-O'Leary, 2005). A third account subsumes the previously mentioned accounts, contending that the general degree of similarity between attributes, as determined by any one of many salient factors, leads to the behavioral differences in the IAT (De Houwer *et al.*, 2005). A fourth theory maintains that the IAT reflects differential costs of task switching, such that costs are smaller in compatible blocks compared with incompatible blocks (Mierke and Klauer, 2003).

While these theories present sound arguments for the types of cognitive processing that underlie IAT effects, it is important to note that they are based purely on behavioral measures and findings. One way to gain more insight into the ongoing cognitive processes evident during the IAT is to use neuroelectric measurement in conjunction with behavioral measures. Neuroelectric activity occurs continuously during the completion of a task and the temporal sensitivity of neuroelectric measurement allows for investigation into the subset of cognitive processes that occur before, during and after stimulus encoding and response production, which may not be manifest at discrete behavioral levels. As such these measures, which include event-related brain potentials (ERPs), should enable researchers to more precisely examine the cognitive processing that occurs during an IAT relative to behavioral measures; not only in relation to the types or qualities of the processes, but also in relation to the timing of the processes and when, if ever, they start to differ.

Currently, few have studied the neural bases for typical findings in IATs or for implicit measures in general. Stanley *et al.* (2008) did present a review of neurological research examining implicit processing; however, the review consisted mainly of fMRI and PET studies that indicated which brain areas were involved in processing, but little about the types of processing or the timing of that processing. Hurtado and colleagues (Hurtado *et al.*, 2009) examined ERPs during the

Received 25 September 2009; Accepted 8 June 2010
Advance Access publication 2 July 2010

Correspondence should be addressed to John K. Williams, Rosemead School of Psychology, 13800 Biola Avenue, La Mirada, CA 90639, USA. E-mail: john.williams@biola.edu

execution of an IAT and found effects in the late positive potential (LPP); however, the LPP was the only component analyzed. O'Toole and Barnes-Holmes (2009) also examined ERPs during IAT execution and found several effects, but only focused on later ERP components (300+ ms), did not use a group bias IAT, and limited their discussion of the IAT to comparisons with semantic priming tasks. Finally, He and colleagues (He *et al.*, 2009) examined multiple ERPs across a wide range of temporal latencies and found both early (P2, N2) and late (LPC) correlates to IAT behavioral effects, but these ERPs were measured during a racial gender identification task, not during the IAT.

Given the paucity of this literature and the varied findings and methodologies utilized, more research is needed to better understand the ongoing nature of the cognitive processes present during IAT execution. Accordingly, the present study is exploratory in nature; it examines the effects of the compatible and incompatible conditions of a group bias IAT on numerous early and late ERP components, including the N1, P2, N2, N400 and LPP, and their relationship to response time differences. The multiple components are examined to not only build upon initial examinations of the IAT and ERPs described above, but also to better determine both the type of processing differences that may exist (e.g. attentional, perceptual, response conflict, semantic and emotional, respectively) and the timing of those processing differences (i.e. whether differences occur early or late in the overall processing).

Components to be measured

The N1 is thought to reflect selective attention differences to stimulus characteristics and intentional discrimination processing (Vogel and Luck, 2000) with increased N1 amplitudes reflecting enhanced processing of attended stimuli (Luck, 1995), and no effects due to arousal (Vogel and Luck, 2000) or inhibition (Bokura *et al.*, 2001). Accordingly, N1 differences in the IAT would suggest selective attentional differences may exist in the processing of compatible and incompatible stimuli.

The P2 is generally believed to be associated with perceptual processing of stimuli (Doyle *et al.*, 1996). However, ERP priming and implicit memory effects have also been evident in P2 amplitudes (Rugg *et al.*, 1994) and research has shown P2 differences related to directing attention toward negative information relative to positive information (Bartholow and Dickter, 2007). Therefore, if perceptual processing, implicit memory or information valence differences exist across compatibility in the IAT, the amplitude of the P2 component should be sensitive to these differences.

The N2 has multiple psychological interpretations, but one that is particularly relevant to the IAT is the interpretation that the N2 reflects the amount of response conflict present during task execution (Yeung *et al.*, 2004). Specifically, the amplitude of the N2 is thought to reflect the amount of conflict present between multiple response options immediately prior to selecting and executing the appropriate motor response to correctly complete the task. In relation to the IAT, it follows that greater levels of response conflict would be present during incompatible trials compared to compatible trials, which would be indicated by larger (more negative) N2 amplitudes.

The N400 has been widely used as a measure of semantic congruency in both sentences (Kutas and Hillyard, 1980a, b) and individual words (Holcomb, 1988; Kiefer, 2002). In tasks where semantically incongruent word pairs or sentences are presented (e.g. 'The boy ate the couch.'), N400 amplitudes are larger than when participants view comparable semantically congruent statements, suggesting N400 amplitude is sensitive to semantic violations. Thus it is possible that N400 amplitudes will be larger during incompatible trials of the IAT.

The LPP has been used in psychophysiological research as an indicator of several different cognitive processes (Cacioppo *et al.*, 1994; Dillon *et al.*, 2006; Hajcak and Nieuwenhuis, 2006). LPP amplitudes are found to increase when emotional stimuli are being processed, with larger amplitudes for negative stimuli *vs* positive stimuli, and increased amplitudes for positive stimuli compared to neutral stimuli (Cuthbert *et al.*, 2000; Hajcak and Nieuwenhuis, 2006; Hajcak *et al.*, 2006; Zilber *et al.*, 2007). Additionally, LPP amplitudes are enhanced when the emotional stimulus is made more salient, seemingly by increasing attentional resources recruited for the task. This increased saliency can be accomplished by improving the semantic cohesion of stimuli being presented (Dillon *et al.*, 2006), or by pairing together stimuli in different modalities (e.g. pictures and words, pictures and sounds) with similar affective valence (Spreckelmeyer *et al.*, 2006). In both cases, LPP amplitudes can be enhanced by increasing the emotional congruency of stimuli presented close together in time, and it seems likely that processing of compatible stimuli in a group bias IAT would lead to larger LPP amplitudes.

The present study attempts to illuminate the early and late cognitive processes that lead to RT differences across conditions of a gay-straight IAT. Given the previous findings from IATs measuring implicit attitudes toward homosexuals (Banse *et al.*, 2001; Steffens and Buchner, 2003; Boysen *et al.*, 2006), it is predicted that RTs will be faster in the compatible conditions (i.e. gay-negative and straight-positive pairings) than in the incompatible conditions (i.e. straight-negative and gay-positive pairings). Also, based on the ERP literature presented above and the nature of the present IAT methodology, multiple ERP predictions are present. It is predicted that the early components (N1, P2) will not differ across IAT conditions as the present task conditions do not vary in relation to stimulus characteristics, selectivity of attention, perceptual processing or the amount of negative relative to positive information presented. Conversely, it is predicted that later ERP effects will be evident during the IAT, with smaller N2 and N400 amplitudes and larger LPP amplitudes in compatible conditions compared to incompatible conditions. These findings would suggest less response conflict (N2), and both greater semantic (N400) and emotional (LPP) congruency in the compatible condition compared to the incompatible condition. Finally, it is predicted that the degree of behavioral (RT) difference across the two IAT task conditions will be related to the degree of neural activity difference across task conditions as evidenced by differences in N2, N400 and LPP amplitudes.

METHODS

Participants

A total of 23 undergraduate students (16 females and 7 males) from a Christian Liberal Arts University in Southern California volunteered for the experiment. All participants were given extra course credit for their participation. Data from one participant was excluded from analysis due to outlying IAT response times (>2 s.d. from the mean) and two additional participants were excluded due to excessive noise and artifacts in their EEG recordings.

IAT stimuli and procedure

Similar to the gay-straight IAT used by Boysen *et al.* (2006), gay targets were represented by eight photographs of either two men or two women, while straight targets were represented by eight photographs of a man and a woman. Good and bad attributes were represented by eight positive or eight negative adjectives used previously by Greenwald *et al.*, (1998). Participants were instructed to press one of two keys to classify targets or attributes, and were given feedback via a red ‘X’ in the middle of the screen if an incorrect response was recorded; these responses were not required to be corrected. The inter-trial interval was 1000 ms for all trials.

Each participant completed two 16-trial practice blocks. One block was used to learn to classify good and bad attributes by pressing one of two response keys and the other to learn the gay-straight target distinction. Identification of the targets and attributes were then combined in a single block of 32 trials in which one attribute and target were paired together on one response key (e.g. positive-gay), while the other attribute–target pair (e.g. negative-straight) was assigned to the other response key. During this combined practice block, labels (e.g. ‘good-gay’ or ‘bad-straight’) were visible in the upper left and right hand corners of the screen to remind participants which button to press for each stimulus. In order to reduce eye movements in the test blocks, participants then completed a second combined practice block of 32 trials with the labels removed from the screen to prepare for the subsequent ERP data collection. Given the need to collect enough data for ERP analysis, four test blocks of 64 combined trials were then completed while EEG data was recorded. Participants were given a rest period between each block of self-determined length, and the entire experiment lasted between 30 and 45 min, depending on the length of the rest periods.

Based on previous findings (Boysen *et al.*, 2006), it was assumed that the ‘gay-positive’ and ‘straight-negative’ pairings would constitute the ‘incompatible’ condition, while the ‘gay-negative’ and ‘straight-positive’ pairings would represent the ‘compatible’ condition. All four practice blocks and four test blocks were then repeated with the opposite pairings of attributes and targets. The order of the compatible and incompatible blocks varied across participants. One group of participants performed all of the compatible blocks first, another group performed all of the incompatible blocks first, and a final group alternated between compatible and incompatible blocks throughout their participation (initial conditions were counterbalanced across participants in this group as well).

Neural assessment

The electroencephalogram (EEG) was recorded from 64 sintered Ag–AgCl electrodes embedded in an elastic cap, arranged in an extended 10–20 system montage with a ground electrode (AFz) on the forehead. The sites were referenced online to a midline electrode placed at the midpoint between Cz and CPz. Vertical and horizontal bipolar electrooculographic activity (EOG) was recorded to monitor eye movements using sintered Ag–AgCl electrodes placed above and below the right orbit and near the outer canthus of each eye. Impedances were kept below 10 kΩ for all electrodes. A Neuroscan Synamps2 bioamplifier (Neuro Inc., El Paso, TX, USA), with a 24 bit A/D converter and $\pm$200 millivolt (mV) input range, was used to continuously digitize (1000 Hz sampling rate), amplify (gain of 10) and filter (70 Hz low-pass filter, including a 60 Hz notch filter) the raw EEG signal in DC mode (763 μV/bit resolution). EEG activity was recorded using Neuroscan Scan software (v 4.3.1). Stimulus presentation, timing, and measurement of behavioral response time and accuracy were controlled by E-Prime (v 2.0) software.

Offline neural processing of the stimulus-locked components included: re-referencing to average mastoids, creation of stimulus-locked epochs (from -100 to 1000 ms relative to stimulus presentation), baseline removal (100 ms time window running from -100 to 0 ms prior to the stimulus), low-pass filtering (30 Hz; 24 dB/octave) and artifact rejection (all epochs with signal that exceeded $\pm$100 μV, including eye movements and blinks, were rejected). Separate average ERP waveforms were created for each task stimulus (picture, word) and condition within the IAT paradigm (compatible, incompatible), with an average number of 192 trials (s.d. = 47) for the compatible condition and 189 trials (s.d. = 41) for the incompatible condition. All ERP components were quantified as the average amplitude in a discrete time window following stimulus presentation in each of the two average waveforms for each stimulus. Specifically, N1 was quantified post-stimulus from 110–170 ms for both pictures and words, P2 from 170–240 ms (pictures) and 170–300 ms (words), N2 from 240–370 ms for pictures, N400 from 400–520 ms (pictures) and 300–460 (words), and LPP from 530–1000 ms (pictures) and 500–1000 ms (words) in each of the stimulus-locked average waveforms (Figure 1).

Statistical analyses

Statistical analyses were performed separately for each ERP component. All mean amplitude ERP values were submitted to a 2 (stimulus: picture, word) × 2 (compatibility: compatible, incompatible) × 6 (site: Fz, FCz, Cz, CPz, Pz, Oz) × 3 (order: all compatible first, all incompatible first, mixed blocks) mixed-model ANOVA. Behavioral data (RT) were submitted to a 2 (stimulus: picture, word) × 2 (compatibility: compatible, incompatible) × 3 (order: all compatible first, all incompatible first, mixed blocks) mixed model ANOVA. All analyses with three or more within-subjects levels used the Wilks' lambda statistic. The alpha level for all tests was set at $P = 0.05$. Since the focus of this study was on the compatibility effects of the IAT, only those effects that are related to compatibility are discussed in the main text, while a more detailed account stimulus and site effects are reported in the Supplementary Data.

RESULTS

IAT

On average, participants performed at 95% accuracy in the compatible condition and 92% accuracy in the incompatible condition of the IAT, with all participants performing above 80% accuracy in each condition. Erroneous responses were excluded from the analyses. For each participant, mean response latencies (RTs) were calculated for incompatible (gay-positive, straight-negative) and compatible (gay-negative, straight-positive) trials, as well as for pictures (gay, straight) or words (good, bad). The omnibus analysis revealed a main effect for compatibility, $F(1,17) = 34.7$, $P < 0.001$, partial $\eta^2 = 0.64$, with significantly longer RTs for incompatible trials ($M = 729.2$ ms, s.e. = 27.2) compared with compatible trials ($M = 644.0$ ms, s.e. = 18.5), confirming the expected IAT effect. No other significant effects were observed.

N1

The omnibus analysis for the N1 revealed a main effect for site, $F(5,13) = 5.9$, $P = 0.005$, partial $\eta^2 = 0.69$, with larger amplitudes frontally and centrally, and smaller amplitudes at more posterior sites. Also significant was the effect of stimulus type, $F(1,17) = 34.36$, $P < 0.001$, partial $\eta^2 = 0.67$, with larger amplitudes for pictures than words. No other significant effects were observed, including effects of compatibility, suggesting no attentional differences across compatible and incompatible trials in this IAT task. Based on visual inspection of the data, a subsequent condition × order analysis of the N1 was conducted using only amplitude scores at FCz. No significant effects were observed.

P2

The omnibus analysis for P2 revealed a significant effect of stimulus, $F(1,17) = 18.98$, $P < 0.001$, partial $\eta^2 = 0.53$, with larger amplitudes for words than pictures. In addition, a significant stimulus × site interaction was found, $F(5,13) = 12.44$, $P < 0.001$, partial $\eta^2 = 0.83$, with a greater site effect for pictures than words. No other significant effects were found, suggesting no perceptual or information valence differences across compatibility in this IAT task. Based on visual inspection of the data, a subsequent compatibility × order analyses of the P2 was conducted using only amplitude scores at FCz. Again, no significant effects were observed.

N2

Since the morphology of the waveforms for pictures and words varied, an N2 component was present for pictures but not words (Figure 1). Therefore, the omnibus ANOVA for the N2 was performed only for picture stimuli, and did not include an analysis of stimulus effects. Contrary to our hypotheses, the omnibus N2 analysis revealed no effects of compatibility, suggesting no significant difference in response conflict across compatible and incompatible trials in this version of the IAT. The analysis did reveal a main effect for site, $F(5,13) = 11.6$, $P < 0.001$, partial $\eta^2 = 0.82$, with larger amplitudes frontally and centrally, and smaller amplitudes at more posterior sites. No other significant effects were observed. Based on the extensive response conflict literature (Nieuwenhuis *et al.*, 2003; Yeung *et al.*, 2004) and visual inspection of the data, a subsequent compatibility × order analysis of the N2 was conducted using only amplitude scores at FCz. No significant effects were observed.

N400

As hypothesized, the omnibus analysis for the N400 revealed a main effect for compatibility, $F(1,17) = 8.0$, $P = 0.012$, partial $\eta^2 = 0.32$, with larger amplitudes in the incompatible ($M = -0.97\,\mu V$, s.e. = 0.48) compared to the compatible ($M = 0.11\,\mu V$, s.e. = 0.55) condition, suggesting greater semantic congruency in the compatible condition of the IAT. Further, a site effect was observed, $F(5,13) = 12.90$, $P < 0.001$, partial $\eta^2 = 0.83$, with larger amplitudes frontally and smaller amplitudes at more central and posterior sites. No other significant effects were present. All subsequent analyses of the N400 used amplitude scores at FCz based on the literature (Deveney and Pizzagalli, 2008) and visual inspection of the grand-averaged ERP waveforms.[1]

N400 and IAT

To further examine the relationship between behavioral IAT effects and N400 amplitudes, correlations were run

[1]N400 analyses were also conducted at Cz, CPz and Pz sites. Findings did not differ from those reported in the text at FCz. These analyses were not included in the text to clarify the presentation of the data.

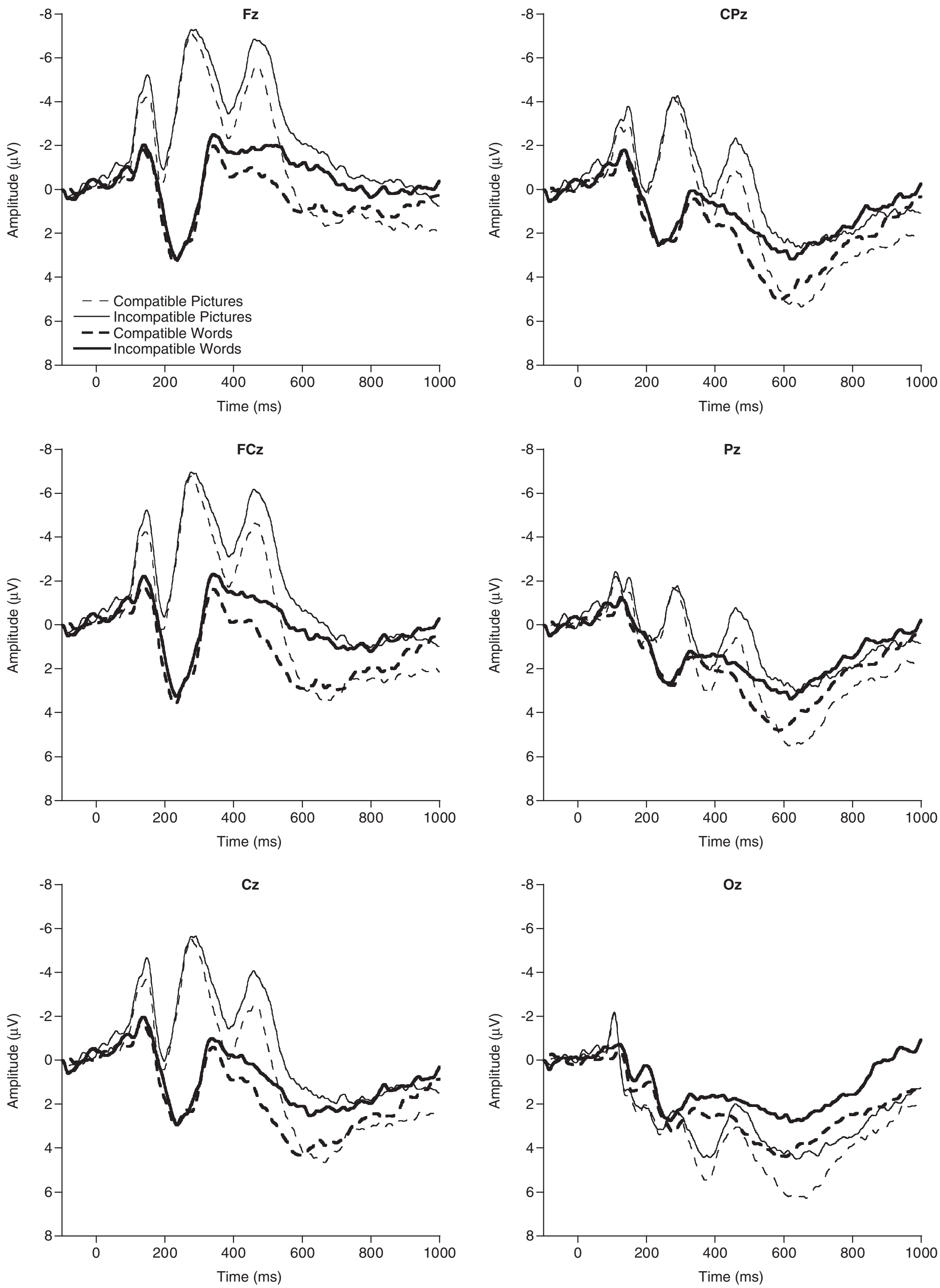

Fig. 1 Grand averaged stimulus-locked waveforms for each stimulus type (pictures, words) and IAT task condition (compatible, incompatible) at the Fz, FCz, Cz, CPz, Pz and Oz electrode sites.

between RTs and N400 amplitudes in each compatibility of the IAT (collapsed across stimulus type), as well as the RT difference (i.e. IAT effect score; incompatible RT–compatible RT) and N400 amplitude difference across IAT compatibility (compatible–incompatible). Table 1 provides the resultant correlation matrix. Significant correlations were present between the IAT effect score and both compatible and incompatible N400 amplitudes, with greater N400 amplitudes associated with larger IAT effect scores, but no relationship between RT differences and N400 amplitude differences. However, visual inspection of the data suggested a curvilinear 'inverted-U' relationship may exist across task conditions (Figure 2a for the scatter plot and quadratic fit estimation). A quadratic analysis was conducted regressing the N400 amplitude difference on the IAT effect score. The overall regression model was significant, $R^2 = 0.31$, $F(2,17) = 3.9$, $P = 0.041$, suggesting that the lowest and highest IAT effect scores in the distribution were associated with smaller or negative differences in N400 amplitude across compatibility while the mid-range of IAT effect differences were associated with larger positive differences in N400 amplitude across compatibility. Table 2a summarizes the regression analysis.

LPP

As hypothesized, the omnibus analysis for the LPP revealed a significant effect for compatibility, $F(1,17) = 16.1$, $P = 0.001$, partial $\eta^2 = 0.47$, with larger (more positive) LPP amplitudes across sites in the compatible ($M = 2.6\,\mu V$, s.e. $= 0.6$) compared with the incompatible ($M = 1.0\,\mu V$, s.e. $= 0.6$) condition, suggesting increased emotional congruency in the compatible condition. However, this main effect was modified by a two-way interaction of compatibility × order, $F(2,17) = 4.6$, $P = 0.026$, partial $\eta^2 = 0.35$. Decomposition of the interaction into *post hoc* Bonferroni-corrected simple compatibility effects for each order indicated that participants who received incompatible blocks of the IAT first showed a significant difference in LPP amplitude, $t(4) = 4.5$, $P = 0.01$, across compatible ($M = 2.6\,\mu V$, s.e. $= 1.1$) and incompatible ($M = -0.3\,\mu V$, s.e. $= 1.4$) conditions of the IAT. No such effect was observed for participants who received compatible blocks first, $t(7) = 1.6$, $P = 0.14$ (compatible $M = 3.7\,\mu V$, s.e. $= 0.9$; incompatible $M = 2.1\,\mu V$, s.e. $= 0.7$), or a mixed order of IAT task blocks during their participation, $t(6) = -0.90$, $P = 0.40$ (compatible $M = 2.7\,\mu V$, s.e. $= 0.8$; incompatible $M = 3.0\,\mu V$, s.d. $= 1.1$).

Further, the analysis revealed a main effect for site, $F(5,13) = 3.6$, $P = 0.029$, partial $\eta^2 = 0.58$, with smaller amplitudes frontally and larger LPP amplitudes at more central

Table 1 Correlations among measures of LPP amplitude and behavioral RTs, LPP amplitudes and N400 amplitudes for each condition of the IAT and for the differences across IAT conditions

Variable	1	2	3	4	5	6	7	8	9
1. IAT RT *C*	–								
2. IAT RT I	0.88*	–							
3. IAT RT *I*—*C*	0.11	0.58*	–						
4. N400 Amp. *C*	−0.24	−0.42	−0.46*	–					
5. N400 Amp. *I*	−0.18	−0.37	−0.46*	0.69*	–				
6. N400 Amp. *C*—*I*	−0.10	−0.10	−0.04	0.47*	−0.32	–			
7. LPP Amp. *C*	−0.01	−0.22	−0.46*	0.60*	0.37	0.34	–		
8. LPP Amp. *I*	−0.12	−0.27	−0.35	0.29	0.66	−0.42	0.60*	–	
9. LPP Amp. *C*—*I*	0.14	0.10	−0.03	0.25	−0.43	0.84*	0.26	−0.62*	–

IAT, implicit association test; RT, response time; *C*, compatible condition; *I*, incompatible condition; *I*—*C*, incompatible minus compatible; Amp., amplitude; *C*—*I*, compatible minus incompatible. *$P < 0.05$.

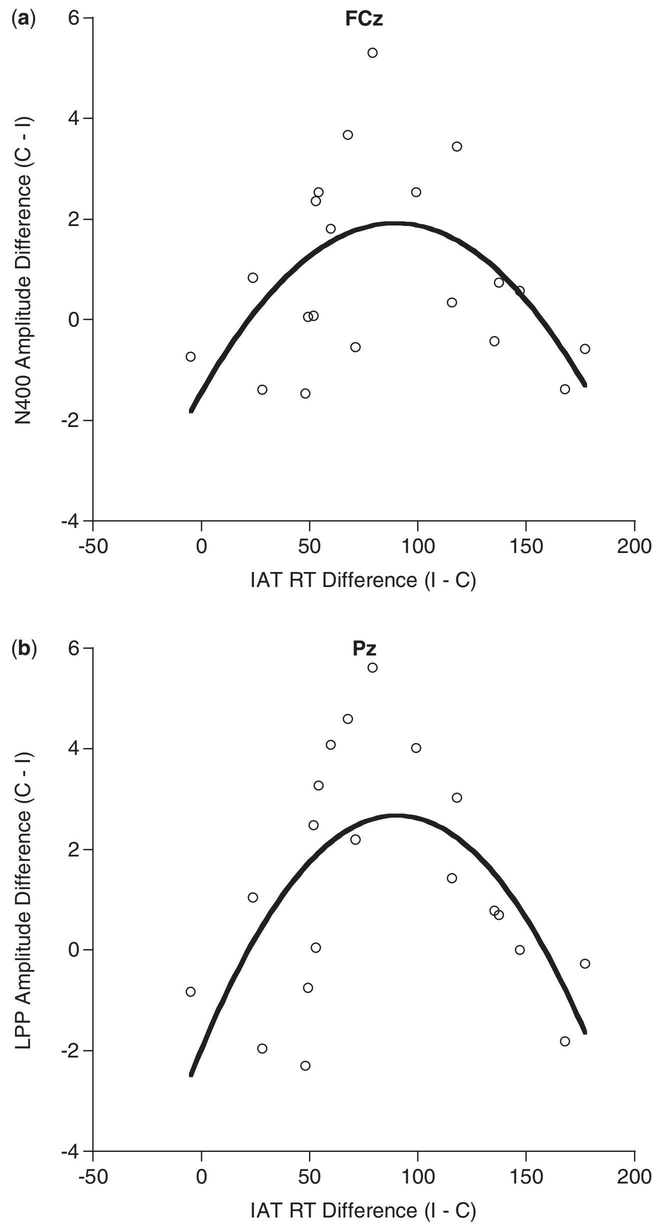

Fig. 2 Scatter plots for the relationship between (**a**) IAT RT differences and N400 amplitude differences at FCz and (**b**) IAT RT differences and LPP amplitude differences at Pz, across task conditions (*C*, compatible; *I*, incompatible) of the IAT.

Table 2 Summary of quadratic regression analyses for variables predicting the difference in N400 amplitudes across conditions of the IAT

Variable	B	s.e. B	β
IAT RT *I—C*	0.08	0.02	2.0*
IAT RT $(I—C)^2$	0.00	0.00	−2.1*

IAT, implicit association test; RT, response time; *C*, compatible condition; *I*, incompatible condition; *I—C*, incompatible minus compatible; s.e.B, Standard Error for B. *$P < 0.05$.

Table 3 Summary of quadratic regression analyses for variables predicting the difference in LPP amplitudes across conditions of the IAT

Variable	B	s.e. B	β
IAT RT *I—C*	0.10	0.03	2.2*
IAT RT $(I—C)^2$	0.00	0.00	−2.4*

IAT, implicit association test; RT, response time; *C*, compatible condition; *I*, incompatible condition; *I—C*, incompatible minus compatible; s.e.B, Standard Error for B. *$P < 0.05$.

and posterior sites. No other significant effects were observed. All subsequent analyses of the LPP used amplitude scores at Pz (Cacioppo *et al.*, 1994; Ito and Cacioppo, 2000; Hajcak and Nieuwenhus, 2006; Moser *et al.*, 2006).[2]

LPP and IAT

Correlations between RT and LPP amplitude in each compatibility (collapsed across stimulus type), as well as the RT difference (i.e. IAT effect score) and LPP difference across IAT compatibility revealed a significant relationship between LPP amplitude in the compatible condition with the IAT effect score, with greater compatible LPP amplitudes associated with smaller IAT effect scores (Table 1), but no relationship between RT differences and LPP amplitude differences. However, visual inspection of the data again suggested a curvilinear 'inverted-U' relationship between the IAT effect score and LPP amplitude differences, similar to the relationship between the IAT and N400 (Figure 2b for the scatter plot and quadratic fit estimation). A quadratic analysis was conducted regressing the LPP amplitude difference on the IAT effect score. The overall regression model was significant, $R^2 = 0.39$, $F(2,17) = 5.4$, $P = 0.02$, suggesting that the lowest and highest IAT effect scores were associated with smaller or negative differences in LPP amplitude across compatibility while the mid-range of IAT effect differences were associated with larger positive differences in LPP amplitude across compatibility. Table 3 summarizes the regression analysis.

DISCUSSION

The purpose of the experiment was to employ ERP measures to examine the neural bases of response time differences in a typical IAT involving group bias. An IAT measuring attitudes toward homosexuals was used in the present study. With regard to the behavioral data, results of the gay-straight IAT were as predicted and mirrored previous findings (Banse *et al.*, 2001; Steffens and Buchner, 2003; Boysen *et al.*, 2006); that is, RTs for the compatible responses (gay-negative, straight-positive) were significantly faster than for the incompatible responses (gay-positive, straight-negative).

In terms of the ERP measures, as expected, the earliest components (N1, P2) showed no differences across conditions of the IAT, suggesting earlier attentional and perceptual processes are not associated with behavioral IAT effects. However, later components (N400, LPP) were modulated by stimulus conditions. More specifically, in the case of the N400, amplitudes were larger (more negative) for incompatible than for compatible trials. This finding is consistent with previous research which implicates the N400 as an indicator of semantic incongruency (Kutas and Hillyard, 1980a, b) and suggests greater semantic incongruity exists for incompatible trials of the IAT (O'Toole and Barnes-Holmes, 2009). Thus, when the concept of gay people and positive attributes are paired together on the same response button, the semantic incongruity between the two appears to lead to a slowing of the response.

With regards to the LPP, amplitudes were larger (more positive) during compatible trials than during incompatible trials for those participants receiving incompatible trials first, mirroring previous studies which show an association between LPP amplitude and the emotional congruency of targets (Dillon *et al.*, 2006; Spreckelmeyer *et al.*, 2006). Therefore, the pairing together of emotionally compatible stimuli on the same response button appears to enhance the processing of the items, especially after receiving incompatible stimulus pairings. This finding expands on previous IAT research in which LPP differences were found at frontal (Hurtado *et al.*, 2009) and central sites (O'Toole and Barnes-Holmes, 2009; though O'Toole and Barnes-Holmes found a reversal at prefrontal sites) suggesting processes in addition to semantic priming, including evaluative processes, are present during IAT execution.

It follows that targets and attributes that are not closely linked, as in the case of incompatible trials, would also be associated with slower RTs during IAT execution. Whether this 'linkage' is rooted in semantics, salience, or general similarity is still debated among theorists. The current N400 findings do not address which theoretical explanation is most viable, as the spreading activation (Greenwald *et al.*, 1998), salience asymmetry (Rothermund and Wentura, 2004), and similarity (De Houwer *et al.*, 2005) accounts are all supported by N400 differences. What the present data reveal, however, is that both semantic and affective properties of the stimuli seem to contribute to the stronger association of the compatible items.

[2]Additional LPP analyses were conducted at FCz, Cz and CPz sites. The pattern of findings did not differ from those reported in the text at Pz. These analyses were not included in the text to clarify the presentation of the data.

Given that the findings showed comparable effects of target-attribute compatibility on both RTs and ERP (N400, LPP) amplitudes, it seemed likely to find a significant relationship between RT and ERP differences across conditions in the IAT. More precisely, it seemed logical to predict a linear relationship between the two, with ERP amplitude differences increasing linearly with IAT RT differences. However, while a significant relationship was found, it was curvilinear in nature. For both the N400 and LPP measures, amplitude differences increased with RT differences for participants who had smaller IAT RT differences. However, participants with larger IAT RT differences actually showed a decrease in amplitude differences as RT differences increased. This suggests that as a participant's bias increases from no bias to a moderate level of bias (as indicated by IAT RT differences), there is an increasing level of semantic and emotional congruency between the targets and attributes in the compatible condition as compared to the incompatible condition. However, as bias increases from moderate to high, there appears to be a reduction of semantic and emotional congruency for the compatible trials as compared to the incompatible trials as evidenced by the smaller difference in LPP amplitudes across conditions. Presently, there is no apparent theoretical explanation for the curvilinear relationship between RT differences and ERP amplitude differences. What is clear from these data is that the cognitive processes involved in a typical IAT task are multifaceted.

While results from the N400 and LPP data were in line with our hypotheses, contrary to our predictions, N2 amplitudes did not vary across the IAT. This finding is inconsistent with the notion that greater response conflict (Yeung *et al.*, 2004) is present during incompatible conditions of the IAT. While this does not suggest that executive control processes or response-related processes are not meaningful during IAT execution, it suggests that competition among potential responses may not be the specific executive process related to alterations in IAT behavior. There is the possibility that executive processes may be acting immediately prior to the response, as evidenced by the flattening in the incompatible waveforms ~600 ms after stimulus presentation at Fz and FCz (Figure 1). It appears, however, that these processes occur later in processing and are likely influenced by the relative semantic and affective congruency of the stimuli. Future research should investigate response-related components to better assess the influence of executive control and pre-response processes on the IAT.

Future directions

Although not a focus of the present investigation, a more thorough examination of order effects in the IAT might prove fruitful given the significant order by compatibility interaction found for the LPP. While the compatibility effect on the LPP was only found for those that completed the incompatible condition first, it is possible that the other orders did not show a similar effect due to small sample sizes. In addition, future research may want to look more closely at the different waveform morphologies produced while processing pictures and words, since IAT research often utilizes varied stimulus presentations. In the current study, the simple stimulus differences were unrelated with compatibility findings, but stimulus type might have an effect in some paradigms. One might also examine more closely the curvilinear relationship found in the present study between RT differences and both N400 and LPP amplitude differences to shed more light on how semantic and emotional congruency influence the processing of IAT targets. Finally, it is possible that the two major ERP findings in the present study may actually be generated from a single component; that is, the N400 may reflect the beginning of the LPP and not a distinct N400. Future research might attempt to look at this possibility more closely.

SUPPLEMENTARY DATA

Supplementary data are available at *SCAN* online.

Conflict of Interest

None declared.

REFERENCES

Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology*, *29*, 369–87.

Banse, R., Seise, J., Zerbes, N. (2001). Implicit attitudes towards homosexuality: reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie*, *48*, 145–60.

Bartholow, B.D., Dickter, C.L. (2007). Social cognitive neuroscience of Person perception: a selective review focused on the event-related brain potential. In: Harmon-Jones, E., Winkielman, P., editors. *Social neuroscience: integrating biological and psychological explanations of social behavior*. New York, NY, US: Guilford Press, pp. 376–400.

Blanton, H., Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist*, *61*, 27–41.

Bokura, H., Yamaguchi, S., Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clinical Neurophysiology*, *112*, 2224–32.

Boysen, G.A., Vogel, D.L., Madon, S. (2006). A public versus private administration of the implicit association test. *European Journal of Social Psychology*, *36*, 845–56.

Cacioppo, J.T., Crites, S.L., Gardner, W.L., Berntson, G.G. (1994). Bioelectrical echoes from evaluative categorizations: I. A late positive brain potential that varies as a function of trait negativity and extremity. *Journal of Personality and Social Psychology*, *67*, 115–25.

Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology*, *52*, 95–111.

De Houwer, J., Geldof, T., De Bruycker, E. (2005). The implicit association test as a general measure of similarity. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *59*, 228–39.

Deveney, C.M., Pizzagalli, D.A. (2008). The cognitive consequences of emotion regulation: an ERP investigation. *Psychophysiology*, *45*, 435–44.

Dillon, D.G., Cooper, J.J., Grent-'t-Jong, T., Woldorff, M.G., LaBar, K.S. (2006). Dissociation of event-related potentials indexing arousal and semantic cohesion during emotional word encoding. *Brain and Cognition*, *62*, 43–57.

Doyle, M.C., Rugg, M.D., Wells, T. (1996). A comparison of the electrophysiological effects of formal and repetition priming. *Psychophysiology*, *33*, 132–47.

Greenwald, A.G., Farnham, S.D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022–38.

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*, 1464–80.

Greenwald, A.G., Nosek, B.A., Sriram, N. (2006). Consequential validity of the implicit association test: comment on Blanton and Jaccard (2006). *American Psychologist*, *61*(1), 56–61.

Hajcak, G., Moser, J.S., Simons, R.F. (2006). Attending to affect: appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion*, *6*, 517–22.

Hajcak, G., Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cognitive, Affective & Behavioral Neuroscience*, *6*, 291–7.

He, Y., Johnson, M.K., Dovidio, J.F., McCarthy, G. (2009). The relation between race-related implicit associations and scalp-recorded neural activity evoked by faces from different races. *Social Neuroscience*, *4*, 426–42.

Holcomb, P.J. (1988). Automatic and attentional processing: an event-related brain potential analysis of semantic priming. *Brain and Language*, *35*, 66–85.

Hurtado, E., Haye, A., González, R., Manes, F., Ibáñez, A. (2009). Contextual blending of ingroup/outgroup face stimuli and word valence: LPP modulation and convergence of measures. *BMC Neuroscience*, *10*. ArtID69.

Ito, T.A., Cacioppo, J.T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology*, *36*, 660–76.

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research*, *13*, 27–39.

Kinoshita, S., Peek-O'Leary, M. (2005). Does the compatibility effect in the race implicit association test reflect familiarity or affect? *Psychonomic Bulletin & Review*, *12*, 442–52.

Kutas, M., Hillyard, S.A. (1980a). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*, 99–116.

Kutas, M., Hillyard, S.A. (1980b). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*, 203–5.

Luck, S.J. (1995). Multiple mechanisms of visual-spatial attention: Recent evidence from human electrophysiology. *Behavioural Brain Research*, *71*, 113–23.

Mierke, J., Klauer, K.C. (2003). Method-specific variance in the implicit association test. *Journal of Personality and Social Psychology*, *85*, 1180–92.

Moser, J.S., Hajcak, G., Bukay, E., Simons, R.F. (2006). Intentional modulation of emotional responding to unpleasant pictures: an ERP study. *Psychophysiology*, *43*, 292–6.

O'Toole, C., Barnes-Holmes, D. (2009). Electrophysiological activity generated during the implicit association test: a study using event-related potentials. *The Psychological Record*, *59*, 207–20.

Rothermund, K., Wentura, D. (2004). Underlying processes in the implicit association test: dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*, 139–65.

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*, 29–43.

Rowatt, W.C., Powers, C., Targhetta, V., Kennedy, S., Labouff, J., Comer, J. (2006). Development and initial validation of an implicit measure of humility relative to arrogance. *The Journal of Positive Psychology*, *1*, 198–211.

Rugg, M.D., Doyle, M.C., Holdstock, J.S. (1994). Modulation of event-related brain potentials by word repetition: Effects of local context. *Psychophysiology*, *31*, 447–59.

Spreckelmeyer, K.N., Kutas, M., Urbach, T.P., Altenmüller, E., Münte, T.F. (2006). Combined perception of emotion in pictures and musical sounds. *Brain Research*, *1070*, 160–70.

Stanley, D., Phelps, E., Banaji, M. (2008). The neural basis of implicit attitudes. *Current Directions in Psychological Science*, *17*, 164–70.

Steffens, M.C., Buchner, A. (2003). Implicit association test: separating transsituationally stable and variable components of attitudes toward gay men. *Experimental Psychology*, *50*, 33–48.

Teachman, B.A., Gregg, A.P., Woody, S.R. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology*, *110*, 226–35.

Vogel, E.K., Luck, S.J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, *37*, 190–203.

White, M.J., White, G.B. (2006). Implicit and explicit occupational gender stereotypes. *Sex Roles*, *55*, 259–66.

Yeung, N., Botvinick, M.M., Cohen, J.D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, *111*, 931–59.

Zilber, A., Goldstein, A., Mikulincer, M. (2007). Adult attachment orientations and the processing of emotional pictures–ERP correlates. *Personality and Individual Differences*, *43*, 1898–907.