



Neural time course and brain sources of facial attractiveness vs. trustworthiness judgment

Manuel G. Calvo^{1,2} · Aida Gutiérrez-García³ · David Beltrán^{1,2}

Published online: 5 September 2018
© Psychonomic Society, Inc. 2018

Abstract

Prior research has shown that the more (or less) attractive a face is judged, the more (or less) trustworthy the person is deemed and that some common neural networks are recruited during facial attractiveness and trustworthiness evaluation. To interpret the relationship between attractiveness and trustworthiness (e.g., whether perception of personal trustworthiness may depend on perception of facial attractiveness), we investigated their relative neural processing time course. An event-related potential (ERP) paradigm was used, with localization of brain sources of the scalp neural activity. Face stimuli with a neutral, angry, happy, or surprised expression were presented in an attractiveness judgment, a trustworthiness judgment, or a control (no explicit social judgment) task. Emotional facial expression processing occurred earlier (N170 and EPN, 150–290 ms post-stimulus onset) than attractiveness and trustworthiness processing (P3b, 400–700 ms). Importantly, right-central ERP (C2, C4, C6) differences reflecting discrimination between “yes” (attractive or trustworthy) and “no” (unattractive or untrustworthy) decisions occurred at least 400 ms earlier for attractiveness than for trustworthiness, in the absence of LRP motor preparation differences. Neural source analysis indicated that facial processing brain networks (e.g., LG, FG, and IPL—extending to pSTS), also right-lateralized, were involved in the discrimination time course differences. This suggests that attractiveness impressions precede and might prime trustworthiness inferences and that the neural time course differences reflect truly facial encoding processes.

Keywords Trustworthiness · Attractiveness · Time course · Event-related potential · Brain sources · Facial expression

Introduction

Human observers very often rely on facial cues to draw social and personality inferences about other people (Sutherland et al., 2015; see Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Particularly, facial appearance is used to readily—from first impressions—infer how trustworthy or

untrustworthy a person may be. More specifically, significant positive *correlations* have been found between perceived facial attractiveness and trustworthiness judgments for faces conveying a neutral expression (Oosterhof & Todorov, 2008; Xu et al., 2012). Furthermore, some emotional facial expressions (happy) are associated with *both* trustworthiness and attractiveness, while others (angry and disgusted) are associated with untrustworthiness and unattractiveness (Sutherland, Young, & Rhodes, 2017; see Said, Haxby, & Todorov, 2011). In addition, some *common* neural networks are recruited during attractiveness and trustworthiness judgments of face stimuli (amygdala, posterior superior temporal sulcus or pSTS, insula, and medial orbitofrontal cortex or mOFC), although there are also unique brain regions for each judgment (see meta-analyses by Bzdok et al., 2011; and Mende-Siedlecki, Said, & Todorov, 2013).

These findings reveal a close relationship between perceived facial attractiveness and trustworthiness. Does this relationship reflect a mere co-occurrence, or does unattractiveness processing occur earlier and therefore can prime and affect un/trustworthiness processing? Presumably,

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13415-018-0634-0>) contains supplementary material, which is available to authorized users.

✉ Manuel G. Calvo
mgcalvo@ull.edu.es

¹ Department of Cognitive Psychology, Campus de Guajara, Universidad de La Laguna, 38205 Tenerife, Spain

² Instituto Universitario de Neurociencia (IUNE), Universidad de La Laguna, 38205 Tenerife, Spain

³ Aida Gutiérrez-García, Department of Health Sciences, Universidad de Burgos, 09001 Burgos, Spain

perception of attractiveness depends on more easily accessible, superficial cues, whereas trustworthiness involves more complex, inner personality traits. It is thus possible that, in the absence of sufficient behavioral evidence (as is the case for most people we encounter in daily life), facial attractiveness is used by observers as a heuristic or mental “shortcut” for inferring trustworthiness. The use of the attractiveness shortcut to trustworthiness would be consistent with a cognitive-economy mechanism, i.e., to minimize processing effort when dealing with information complexity or uncertainty (e.g., when making decisions about trustworthiness if no clear objective data are available). To address this issue, the current study investigated the neural *time course* of facial trustworthiness processing relative to that of attractiveness. If attractiveness is used as a cue and a shortcut for judging trustworthiness, and the former can influence the latter, attractiveness evaluation should, at least, *precede* trustworthiness evaluation.

Event-related potential (ERP) paradigms are well-suited to estimate the time course of neurocognitive processes. To our knowledge, at least 15 published studies using ERP measures have investigated facial attractiveness (e.g., Ma & Hu, 2015; Marzi & Viggiano, 2010; Ohmann, Stahl, Mussweiler, & Kedia, 2016; Schacht, Werheid, & Sommer, 2008; Trujillo, Jankowitsch, & Langlois, 2014; van Hooff, Crawford, & van Vugt, 2011; Werheid, Schacht, & Sommer, 2007; Wiese, Altmann, & Schweinberger, 2014; Thiruchselvam, Harper, & Homer, 2016), whereas at least eight studies have investigated facial trustworthiness (Dzelhyova, Perrett, & Jentsch, 2012; Lischke, Junge, Hamm, & Weymar, 2018; Marzi, Righi, Otonello, Cincotta, & Viggiano, 2014; Ohmann et al., 2016; Rudoy & Paller, 2009; Ruz, Madrid, & Tudela, 2013; Shore, Ng, Bellugi, & Mills, 2017; Yang, Qi, Ding, & Song, 2011). Both facial attractiveness and trustworthiness have been found to modulate ERPs at early and late processing stages, from P1 (~100 ms post-stimulus) to LPP (~400–700 ms) components (the specific findings are reported in the *Discussion* section). However, importantly for this study, in none of these studies were both attractiveness and trustworthiness compared within the same experimental design; rather, ERPs in response to attractiveness or trustworthiness were examined separately. Ohmann et al. (2016) used a common paradigm (in which two faces—rather than one—were presented simultaneously), but the face stimuli were *different* in the attractiveness (natural, real faces) versus the trustworthiness (artificial, computer-generated faces) task. To estimate the *relative* time course of attractiveness and trustworthiness processing—and determine whether the former could prime the latter—a direct comparison is required.

To compare facial attractiveness and trustworthiness processing, we used the following approach. First, *both* attractiveness and trustworthiness judgments were made on the *same* face stimuli, and each judgment was performed by *different* observers to avoid carryover effects across judgment

tasks. Second, in addition to neutral (nonemotional) faces, happy, angry, and surprised faces were included as stimuli to examine whether attractiveness and trustworthiness neural processing interact with (or, rather, generalize across) type of expression. Importantly, this also served to compare the processing time course of emotional expression relative to that of attractiveness and trustworthiness. This adds to prior ERP research on trustworthiness or attractiveness, which has generally used only nonemotional face stimuli. Third, the face stimuli were presented in either an *attractiveness* task or a *trustworthiness* task, where participants judged whether the person in the photograph was attractive or not, or trustworthy or not. In a control condition, participants had to pay attention to the identity of the person in each photograph (“in preparation for a later memory test”). This served to induce active viewing of the faces (as was the case for the other two tasks), while not requiring any explicit judgments (which were instead required for the other tasks). Comparisons of the control condition with the others served to separate the role of task from that of emotional expression.

ERP measures assessed the *time course* of neural activity. First, we focused on (a) N170, which reflects perceptual encoding of faceness (i.e., facial configuration; Rossion & Jacques, 2012); (b) EPN, which involves affective evaluation (Hajcak, Weinberg, MacNamara, & Foti, 2012); and (c) P3b, which is associated with semantic categorization (Polich, 2012). If attractiveness and/or trustworthiness are linked with the configural (perceptual) processing of faceness, they should modulate N170 (~150–200 ms). If they involve emotional processing, EPN (~200–350 ms) modulation will emerge. If effects occur on P3b (~300–600 ms), they will reveal elaborative conceptual processing. Importantly, if attractiveness impressions prime trustworthiness inferences, an *earlier* neural modulation should appear in the attractiveness task. Second, we compared the brain activity preceding Yes versus No (i.e., “attractive” vs. “unattractive”; or “trustworthy” vs. “untrustworthy”) responses. A significant difference in the amount of neural activity between Yes and No decisions will indicate that a viewer has accumulated sufficient evidence in favor of one of them, thus revealing discrimination (e.g., between attractive and *unattractive*; or between trustworthy and *untrustworthy*). We predict an *earlier* neural discrimination for attractiveness than for trustworthiness.

A complementary approach served to indicate whether the predicted neural differences between Yes vs. No decisions reflect truly *encoding* processes of the *face stimulus significance* as un/attractive or un/trustworthy, rather than response preparation processes (i.e., planning and selection of left or right manual pressing). First, by means of the LRP (Lateralized Readiness Potential) (Smulders & Miller, 2012; Fugeras & Naccache, 2016) neural component, we examined the possible contribution of *response preparation* processes. If the predicted earlier differences between Yes and No

responses for attractiveness than for trustworthiness reflect genuine encoding (rather than motor) processes, then the attractiveness temporal advantage will occur in the absence of LRP differences. Second, by means of the LAURA (Local Autoregressive Average) algorithm (Grave de Peralta, González, Lantz, Michel, & Landis, 2001; Michel et al., 2004), we explored the *brain sources* underlying the scalp ERP activity preceding Yes vs. No decisions. If the predicted neural discrimination advantage reflects truly encoding of the face significance, then differences will appear between attractiveness and trustworthiness in brain regions related to face and facial expression processing (e.g., Fusiform Gyrus [FG] or Superior Temporal Sulcus [STS]) (Haxby, Hoffman, & Gobbini, 2000; Wegrzyn et al., 2015), rather than merely in areas related to response preparation processes (e.g., [Supplementary Motor Area \[SMA\]](#), or Premotor Cortex [PRE]).

Methods

Participants

Seventy-five university undergraduates (51 females and 24 males; aged 18–27 years; $M = 22.1$) participated for course credit after informed consent. Twenty-five participants were randomly assigned to an attractiveness judgment task, a trustworthiness judgment task, or a no-judgment control condition, with approximately the same proportion of each sex (16–17 females) in each condition. With this sample size, we had 0.80 power to detect an effect of $f = 0.25$ ($d = 0.5$) at $\alpha = 0.05$, for an F test (repeated-measures ANOVA within-between interaction, with three between-subjects and four within-subjects conditions) (G*Power 3.1; Faul, Erdfelder, Lang, & Buchner, 2007). The study was approved by the local ethics committee (CEIBA: University of La Laguna) and conducted in accordance with the WMA Declaration of Helsinki 2008.

Recent studies investigating ERPs to facial attractiveness or trustworthiness evaluation used sample sizes ranging between 21 (Shore et al., 2017) and 29 (Lischke et al., 2018) participants. Furthermore, in the only study (to our knowledge) that compared a trustworthiness task and an attractiveness task in a between-subjects design (Ohmann et al., 2016)—like the current study—the participant sample ranged between 22 and 24. Accordingly, our choice of 25 valid participants for each task condition was comparable to these prior approaches.

Stimuli

For all three tasks, we used photographs from the Karolinska Directed Emotional Faces (KDEF; Lundqvist, Flykt, & Öhman, 1998; see <http://kdef.se/>) stimulus set. The face

stimuli portrayed 24 individuals (12 females: 01, 07, 09, 11, 13, 14, 19, 20, 22, 26, 29, 33; and 12 males: 03, 05, 06, 10, 11, 12, 18, 22, 23, 24, 25, 35), each posing neutral, happy, angry, and surprised expressions. Neutral faces were included as a nonemotional condition. Happy faces were included because of their emotionally positive valence and because they are typically judged as more attractive and trustworthy than others (Sutherland et al., 2017; see, Said et al., 2011). Angry faces represented negatively valenced expressions and are typically judged as the most unattractive and untrustworthy (Sutherland et al., 2017; Said et al., 2011). Surprised expressions were chosen, because, while being emotional, they are ambiguous (i.e., amenable to both a positive and a negative interpretation), and they are not associated to un/attractiveness or un/trustworthiness (see Said et al., 2011).

In addition, for each face, a specific mask was generated by scrambling the face pixel-by-pixel, with the same low-level image properties. As assessed by Matlab 7.0 (The Mathworks, Natick, MA), mean luminance ($M = 0.289$; $SE = 0.003$), SD luminance ($M = 0.183$; $SE = 0.001$), root mean square contrast ($M = 0.638$; $SE = 0.002$), skewness ($M = -0.491$; $SE = 0.014$), kurtosis ($M = 1.89$; $SE = 0.008$), signal-to-noise ratio ($M = 3.45$; $SE = 0.026$), and entropy ($M = 6.15$; $SE = 0.013$) were identical for the face stimuli and their respective masks (see the [Supplemental Dataset_Low-level Properties of Faces and Masks](#)). Such masks thus provided a baseline condition for all the face stimuli and were presented on each trial immediately before the respective face (see EEG experiment: Procedure and design, below). We aimed to minimize the influence of merely physical properties of the face stimuli on the EEG responses at early stages and therefore to maximize the impact of expressive information.

Norming study: Attractiveness and trustworthiness ratings

Before the ERP experiment, a norming study on the experimental face stimuli was conducted to examine the relationship between attractiveness and trustworthiness, as well as the association of particular expressions to attractiveness and trustworthiness.

A different sample of 48 undergraduates were presented with the 96 face stimuli (24 of each expression: neutral, happy, surprised, and angry) for rating either trustworthiness (24 participants: 12 females) or attractiveness (24 participants: 12 females). Individual face stimuli were presented on a computer screen with E-Prime 2.0 software in four blocks of 24 trials (randomized). Participants judged how attractive the face was on a 1 (“very unattractive”) to 9 (“very attractive”) scale, or how trustworthy on a 1 (“very untrustworthy”) to 9 (“very trustworthy”) scale, by pressing a key out of nine (in the upper row of a computer keyboard). Participants were told that the

study was concerned with first impressions, and they should make their decisions as quickly as possible.

A 2 (Task) \times 4 (Facial Expression) ANOVA was conducted on attractiveness and trustworthiness (1–9) ratings. The raw data are shown in the [Supplemental Dataset_Attractiveness and Trustworthiness Ratings](#). Effects of expression, $F(3, 92) = 103.99$, $p < 0.001$, $\eta_p^2 = 0.77$, appeared. One-way (4: Expression) ANOVAs for trustworthiness, $F(3, 92) = 155.47$, $p < 0.001$, $\eta_p^2 = 0.84$, and attractiveness, $F(3, 92) = 48.36$, $p < 0.001$, $\eta_p^2 = 0.61$, followed by Bonferroni corrections ($p < 0.05$) for post hoc multiple comparisons, revealed that happy faces were judged as both more trustworthy ($M = 6.75$; $SD = 0.50$) and attractive ($M = 5.49$; $SD = 0.47$) than surprised ($M = 4.69$; $SD = 0.66$; and $M = 4.27$; $SD = 0.86$) and neutral ($M = 4.42$; $SD = 0.68$; and $M = 3.98$; $SD = 0.78$) faces, which did not differ from each other and were more trustworthy and attractive than angry faces ($M = 3.26$; $SD = 0.41$; and $M = 3.13$; $SD = 0.46$).

Trustworthiness and attractiveness were significantly correlated for the whole set of stimuli ($N = 96$; $r = 0.89$, $p < 0.001$; 95% CI [0.85, 0.92]) and also for each expression ($n = 24$; all $ps < 0.001$, two-tailed): happiness ($r = 0.73$, [0.46, 0.88]), surprise ($r = 0.72$, [0.45, 0.87]), neutral ($r = 0.72$, [0.45, 0.87]), and anger ($r = 0.69$, [0.40, 0.86]). Figure 1 shows the variability of trustworthiness and attractiveness rating scores for each expression.

EEG experiment: Procedure and design

For the EEG experiment, stimulus presentation on a computer monitor and manual response data collection were controlled by means of Presentation software (version 15.1, Neurobehavioral Systems, Inc.). Each face stimulus was 11.7-cm high by 8.8-cm wide, equalling a visual angle of 9.55° (vertical) \times 7.19° (horizontal) at 70-cm viewing distance. On each trial, after a 500-ms central fixation cross, a mask was presented for 500 ms, followed by a 1-s face display. After face offset, a black screen with the word “Respond” appeared until the participant pressed a “Yes” or a “No” key (with their left or right forefingers), depending on type of task (attractiveness, trustworthiness, or control). “Yes” and “No” response-keys were counterbalanced for left and right hand across participants and tasks. There was a 2-s inter-trial interval.

In the attractiveness, trustworthiness, and control tasks, each participant was presented with 48 trials of each of the 4 expressions (twice each face), randomly, in 192 experimental trials (8 blocks of 24 trials each), following 12 practice trials. Within each block, trial order was randomly established for each participant, and there was the same number of trials of each expression. Judgment performance responses (“Yes” or “No”) and reaction times (time-locked to the offset of the face stimulus) were collected. Participants were asked to withdraw

their manual response until the face disappeared and a verbal label (e.g., “Respond”) appeared.

In the *attractiveness* and the *trustworthiness* tasks, participants judged whether the person in the photograph was attractive or not, or trustworthy or not. In the *control* condition, participants were told that faces would be presented during a “learning” phase, and that their task was to pay attention to them for a later “memory” test of person identities. On each trial in the learning phase (control task), following face offset, the words “Press Left” or “Press Right” (50% of times each) appeared, indicating the key to be pressed. A recognition test was performed at the end of the session, with 24 KDEF faces (50%, new; 50%, old), although measures were not collected.

The experimental design combined a between-subjects factor, Judgment Task (3: attractiveness vs. trustworthiness vs. control), and a within-subjects factor, Facial Expression (4: neutral, happy, angry, and surprised). Twenty-five participants saw all the stimuli (192 trials) in only one task condition. *Task* was a *between-subjects* factor to avoid carryover (exposure, sensitization, habituation, etc.) effects across tasks. The *control* condition was designed to induce active attention to the faces, while not directing viewers to explicitly judge any particular characteristic. This allowed us to separate the role of task, and thus examine implicit, involuntary processing of trustworthiness, attractiveness, and emotional expression.

EEG recording and preprocessing

EEG and EOG signals were recorded using Ag/AgCl electrodes mounted in elastic Quick-caps (Neuromedical Supplies, Compumedics Inc., Charlotte). EOG signal was measured from two bipolar channels: One was formed by two electrodes placed at the outer canthus of each eye; another, by two electrodes below and above the left eye. EEG signal was recorded from 60 electrodes arranged according to the standard 10 to 20 system. All EEG electrodes were referenced online to an electrode at vertex, and recomputed off-line against the average reference. EEG and EOG signals were amplified at 500-Hz sampling rate using Synamp2 amplifier (Neuroscan, Compumedics Inc., Charlotte), with high- and low-pass filter set at 0.05 and 100 Hz, respectively. Electrode impedance was kept below 5 k Ω .

EEG data preprocessing and analysis was conducted using custom scripts in MATLAB (Mathworks, Natick, MA), drawing on Fieldtrip Toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) functions for reading in data files, artifact correction and topographical representations of ERPs. The following transforms were applied to each participant’s dataset. Epochs were extracted with an interval of 700 ms preceding and 2,000 ms following face onset. Trials with drifting or large movement artifacts were removed by visual inspection before correction. Next, independent component analysis (ICA) was applied to remove the effects of blinks

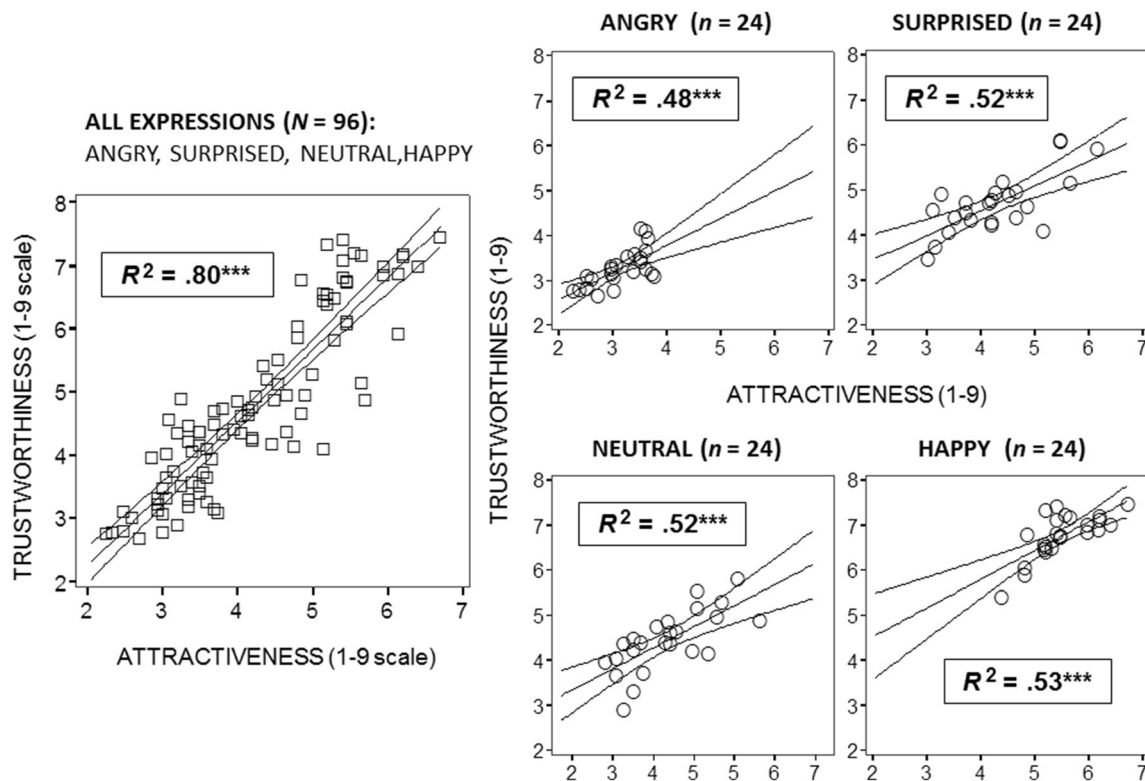


Fig. 1 Attractiveness and trustworthiness judgments. Graphical representation of the relationship between attractiveness and trustworthiness for each expression ($n = 24$ stimuli), with coefficients of

determination (R^2) and prediction of the mean at 95% confidence interval. R^2 : amount of variance in trustworthiness accounted for by attractiveness (or vice versa). *** $p < 0.0001$

and eye movements. Remaining trials with EEG voltages exceeding $70 \mu\text{V}$ measured from peak to peak at any channel were also removed. For ERP computation, artifact-free segments were averaged separately per subject and either facial expression (4) or response (2) conditions (see below). Seven percent of trials were excluded because of artifacts (mainly, eye blinks, drifts, and saccades); ~ 179 valid trials were analyzed per participant (~ 45 per expression condition on each task). Baseline correction of averaged data was performed using the interval between 700 and 500 ms preceding face onset, corresponding to the 200-ms premask onset.

Scalp ERP analysis

Two approaches were used to analyze the ERP data. First, we evaluated the *role of expression and task* in the neural activity of N170, EPN, and P3b components. Second, we examined the *time course of discrimination* of attractive (Yes responses) versus unattractive (No responses) faces, and trustworthy (Yes) versus untrustworthy (No) faces, by means of amplitude-based and onset latency analyses of ERP activity.

Role of expression and task

Three clusters of scalp sites were defined: left parieto-occipital (P5, P7, PO5, PO7, O1), right parieto-occipital (P6, P8, PO6,

PO8, O2), and centro-parietal (CP1, CPZ, CP2, PZ). For each cluster, the mean ERP activity was computed within specific time intervals. For *parieto-occipital* clusters, (a) N170 was assessed between 150 and 200 ms following face onset and (b) early EPN, between 210 and 290 ms. Mixed repeated-measures ANOVAs were conducted, with Task (3) as a between-subjects factor and Facial Expression (4) and Hemisphere (2: left vs. right) as within-subject factors to examine lateralization effects on N170 and EPN. For the *centro-parietal* cluster, the mean amplitude between 400 and 700 ms was calculated to assess P3b. Repeated-measures ANOVAs with Task (3) as a between-subjects factor and Facial Expression (4) as a within-subject factor were conducted on the P3b values.

Time course of discrimination: Amplitude-based analyses

ERPs were computed separately for Yes and No responses, and mixed point-wise ANOVAs with Task (3) as a between-subjects factor and Response (2: Yes vs. No) as a within-subject factor were conducted for every data point for 1 second, from face onset to offset. This represented 60 electrodes by 500 time points and served to detect the scalp sites and period that were sensitive to Task by Response interactions. To control for false positives, an interaction was considered to be reliable only when observed for at least 15 consecutive time

points (≥ 30 ms) and three neighboring channels (De Vega, Beltrán, García-Marco, & Marrero, 2015). The average ERP activity in the reliable space-by-time clusters was analyzed by means of Task by Response ANOVAs, followed by pairwise contrasts between Yes and No responses for each task and one-way (Task) ANOVAs. Finally, differences were computed for each task by subtracting the No from the Yes waveforms to determine the onset of discrimination for *attractiveness* and *trustworthiness*. To this end, the difference (Yes-No) waveforms were submitted to the jackknife procedure described next.

Time course of discrimination: Onset latency analyses

Onset latencies also were analyzed for the stimulus-locked and the response-locked Lateralized Readiness Potential (LRP) component. LRP reflects processes related to response preparation, and its latency often is used to determine the timing of response planning and execution (Smulders & Miller, 2012). For the stimulus-locked LRP (s-LRP), the onset latency is thought to indicate when response-related stimulus processing starts before activation of the response; for response-locked (r-LRP) waveforms, the onset latency reflects when response activation begins. This will be useful to resolve whether potential differences between the current tasks (attractiveness and trustworthiness) are due to the time required to evaluate the faces or, rather, the time devoted to motoric actions (i.e., response preparation and execution). The double subtraction method was applied to compute the s-LRP and r-LRP waveforms for each participant, using the electrodes closer to the hand-specific motor-response areas, i.e., C3 and C4 for the left and the right hemisphere, respectively: The C4 waveform was subtracted from the C3 waveform, and the resulting difference waveform for the No response was subtracted from that for the Yes response (Smulders & Miller, 2012). This double subtraction removes general and hemisphere-specific neural activity that is unrelated to response selection and decision.

The onset latency for s-LRP, r-LRP, and the difference waveforms were estimated as follows (Miller, Patterson, & Ulrich, 1998). To control for the influence of high frequency noise on ERP latencies, waveforms were first low-pass filtered at 10 Hz. Then, the filtered waveforms were submitted to a jackknife procedure in which each participant waveform was replaced with that obtained from averaging the waveforms of all the other participants (i.e., *leave-one-out* strategy). Next, the maximal peak amplitude in the jackknifed waveforms was identified and, going backward in time, the point at which the 30% of peak value was reached was taken as the onset latency. Finally, using the resulting jackknifed latency values, we estimated the participant onset latency using the method described in Smulders (2010). These latencies were then analyzed by means of one-way (Task) ANOVAs.

Brain source analysis

Periods with a reliable Task by Response interaction in the ERP surface analyses were selected to explore the brain sources underlying discrimination between Yes and No responses for each task. To this end, we used the LAURA approach (Grave de Peralta et al., 2001), implemented in Cartool software (Brunet, Murray, & Michel, 2011). The solution space was calculated on a realistic head model that included 4,026 solution points, defined in regular distances within the gray matter of a standard MRI (Montreal Neurological Institute's average brain). Current density magnitudes (ampere per square millimeter) at each point were calculated per subject and condition (Yes and No responses), and submitted to paired *t* tests. The source maps were estimated for the comparison between Yes and No responses for each task. Only the maps that showed *t* test values ≥ 3 for at least 10 nearby solution points were selected for further analysis. For these reliable *t*-test maps, regions of interest (ROIs) were formed from the solution points that showed the largest differences (as defined by the *t* values), and their density magnitudes were submitted to a mixed Task (3) by Response (2) ANOVA.

Results

Behavioral data: Judgment ratings

A 3 (Task) \times 4 (Facial Expression) ANOVA was conducted on the probability of judging faces as attractive or trustworthy (Yes responses). Response latencies were not analyzed, because participants had to wait for the response screen (to standardize the on-line brain activity assessment period) before they could press a key, and thus such latencies would not represent truly decision times. The mean scores of judgment ratings (and latencies), *SEs*, 95% CIs, and contrasts are shown in Table 1. The raw data are shown in the [Supplemental Dataset Responses, ERPs, and Sources](#). For all the following analyses, Greenhouse-Geisser corrections for unequal variances were performed, as well as Bonferroni corrections ($p < 0.05$) for post hoc multiple comparisons.

The ANOVA yielded main effects of task, $F(2, 72) = 17.35$, $p < 0.001$, $\eta_p^2 = 0.33$, and expression, $F(3, 216) = 148.19$, $p < 0.001$, $\eta_p^2 = 0.67$, and an interaction, $F(6, 216) = 39.26$, $p < 0.001$, $\eta_p^2 = 0.52$. Separate one-way (Expression) ANOVAs were conducted for each task.

In the *attractiveness* task, $F(3, 72) = 57.69$, $p < 0.001$, $\eta_p^2 = 0.71$, happy faces were more likely to be judged as attractive than surprised and neutral faces (which did not differ from each other), which were more attractive than angry faces (see the multiple contrasts in Table 1). In the *trustworthiness* task, $F(3, 72) = 96.94$, $p < 0.001$, $\eta_p^2 = 0.80$, there were significant differences among all the expressions, with happy

Table 1 Mean (and SDs) probabilities of responding “Yes” in the attractiveness, the trustworthiness, and the control tasks, and response latencies (RTs, in ms), as a function of facial expressions

Measure	Task	Emotional expression			
		Neutral	Surprise	Anger	Happiness
Response	Attractiveness <i>M</i>	0.302 ^b	0.320 ^b	0.093 ^c	0.718 ^a
	<i>SE</i> and 95% CI	0.033 [0.233, 0.370]	0.035 [0.247, 0.393]	0.022 [0.05, 0.139]	0.050 [0.614, 0.821]
	Trustworthiness <i>M</i>	0.274 ^c	0.402 ^b	0.076 ^d	0.858 ^a
	<i>SE</i> and 95% CI	0.035 [0.203, 0.346]	0.040 [0.318, 0.485]	0.025 [0.024, 0.128]	0.042 [0.771, 0.944]
	Control <i>M</i>	0.486	0.496	0.514	0.506
	<i>SE</i> and 95% CI	0.008 [0.47, 0.50]	0.006 [0.48, 0.51]	0.006 [0.50, 0.53]	0.006 [49, 52]
RTs	Attractiveness <i>M</i>	433	451	430	433
	<i>SE</i> and 95% CI	26 [380, 487]	28 [394, 5028]	28 [373, 487]	28 [376, 490]
	Trustworthiness <i>M</i>	488	508	469	473
	<i>SE</i> and 95% CI	19 [448, 528]	22 [463, 552]	18 [432, 507]	19 [434, 513]
	Control <i>M</i>	509	506	513	508
	<i>SE</i> and 95% CI	16 [477, 541]	16 [473, 539]	17 [478, 548]	16 [475, 542]

Note. For each task, across expressions, means with a different letter (superscript) are significantly different ($p < 0.05$, in multiple contrasts, after Bonferroni corrections). Means sharing a letter or with no superscript are equivalent

faces being the most likely to be judged as trustworthy, followed by surprised faces, and by neutral faces, with angry faces as the most untrustworthy (Table 1). In the *control* task, no significant differences emerged.

Neurophysiological data: N170, EPN, and P3b

For *N170* (150–200 ms), the 3 (Task) \times 4 (Expression) \times 2 (Hemisphere) ANOVA, showed a main effect of hemisphere, $F(1, 72) = 32.30$, $p < 0.001$, $\eta_p^2 = 0.31$, with enhanced negativity in the right ($M = -4.29$; $SE = 0.36$; CI [-5.01, -3.57]) vs. the left ($M = -2.74$; $SE = 0.37$; CI [-3.48, -1.99]) parieto-occipital brain areas. There also was a main expression effect, $F(3, 216) = 12.72$, $p < 0.001$, $\eta_p^2 = 0.15$. Post hoc contrasts for multiple comparisons revealed enhanced negativity for angry ($M = -3.63$; $SE = 0.36$; CI [-4.34, -2.91]), happy ($M = -3.57$; $SE = 0.36$; CI [-4.29, -2.85]), and surprised ($M = -3.76$; $SE = 0.35$; CI [-4.46, -3.06]) expressions—which did not differ from each other—relative to neutral faces ($M = -3.10$; $SE = 0.32$; CI [-3.75, -2.46]). This indicates that *N170* was sensitive to emotional expression (across all three task conditions) but not to task ($F < 1$) (Fig. 2).

For *EPN* (210–290 ms) the Task \times Expression \times Hemisphere ANOVA yielded effects of expression, $F(3, 216) = 19.81$, $p < 0.001$, $\eta_p^2 = 0.22$. Post hoc contrasts showed enhanced negativity for angry ($M = -1.63$; $SE = 0.38$; CI [-2.40, -0.87]), happy ($M = -1.52$; $SE = 0.36$; CI [-2.24, -0.79]), and surprised ($M = -1.19$; $SE = 0.38$; CI [-1.94, -0.44]) expressions—which did not differ from each other—relative to neutral faces ($M = -0.76$; $SE = 0.34$; CI [-1.45, -0.07]). Again, this indicates that *EPN* was sensitive to emotional expression

(across all three task conditions), but not to task, $F(2, 72) = 1.82$, $p = 0.17$, *ns.* (Fig. 2).

In contrast, for *P3b* (400–700 ms), a Task \times Expression ANOVA showed an effect of task, $F(2, 72) = 8.01$, $p < 0.001$, $\eta_p^2 = 0.18$. Post hoc contrasts revealed enhanced positive amplitudes on both the trustworthiness ($M = 3.65$; $SE = 0.35$; CI [2.96, 4.34]) and the attractiveness ($M = 3.24$; $SE = 0.35$; CI [2.55, 3.93]) task conditions relative to the control condition ($M = 1.79$; $SE = 0.35$; CI [1.10, 2.48]) (Fig. 2).

Neural time course of attractiveness and trustworthiness judgments

Amplitude-based analysis

Initially, an exploratory point-wise ANOVA, with Task (3) and Response (2: Yes vs. No) as factors, was conducted on ERP amplitudes throughout the 1-s face display for 60 EEG electrodes (see Scalp ERP analyses, above). Reliable interactive effects emerged (see below) from 600 ms from face onset onwards, over right and left central regions (around C3 and C4 electrodes; Fig. 3A). Accordingly, the average activity in the left (C1, C3, C5) and right (C2, C4, C6) sites was computed and analyzed for the period between 600 and 1,000 ms. An overall ANOVA on ERP amplitudes in this period, with Region (2: left vs. right central), Task (3), and Response (2) as factors, yielded a three-way interaction, $F(2, 72) = 7.02$, $p = 0.002$, $\eta_p^2 = 0.16$, which was decomposed by means of separate Task by Response ANOVAs for each region.

For the *left central* region (C1, C3, C5; see Fig. 3A, left plot, and the right panel), a main effect of response, $F(1, 72) =$

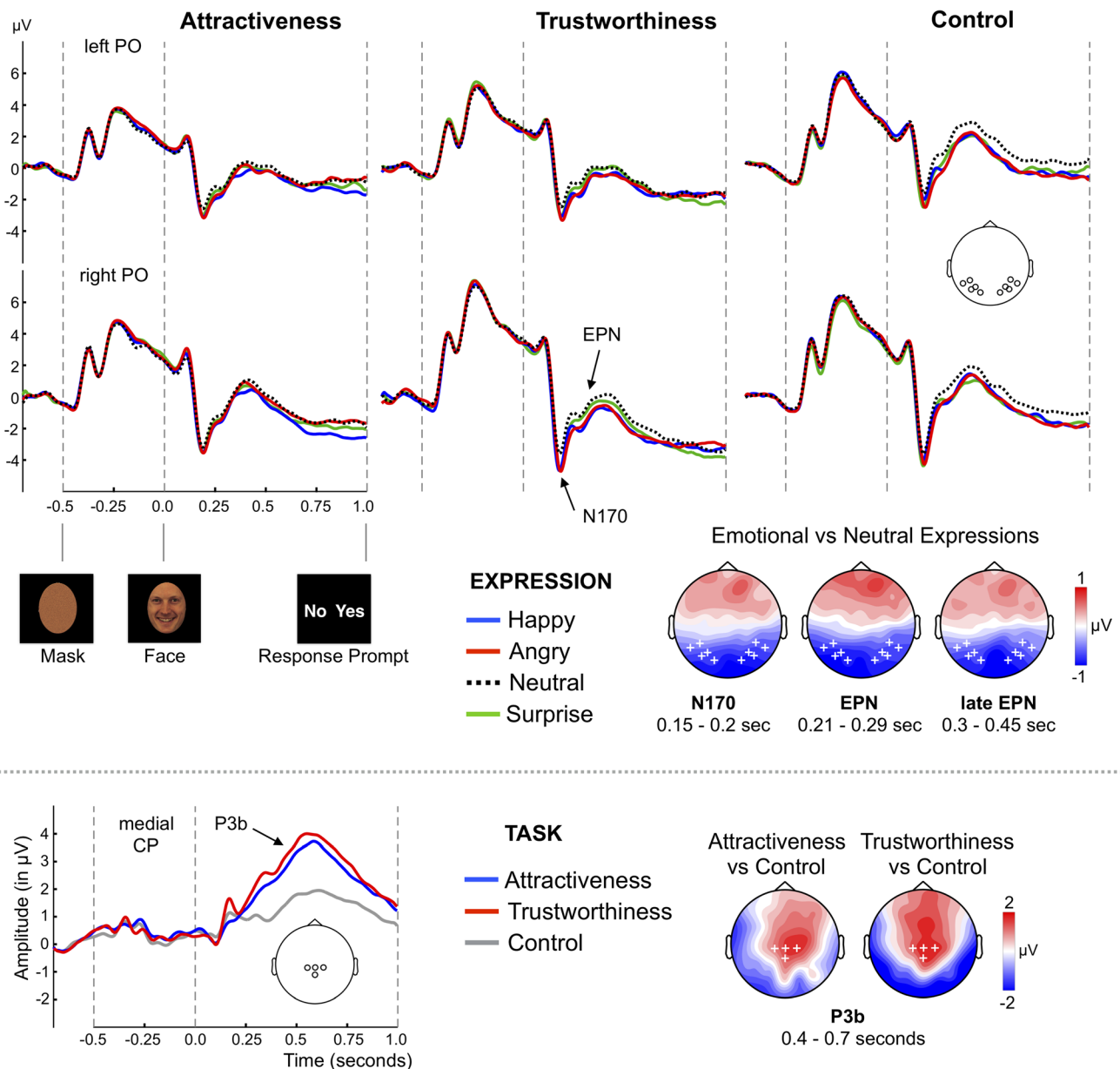


Fig. 2 ERP waveforms and neural maps. *Upper panel:* ERPs elicited at parieto-occipital sites as a function of *expression*. Labels associated with arrows (N170, EPN) indicate ERPs with significant effects. Colored maps represent differences between emotional and neutral expressions. Circles and crosses on maps highlight the sites used to compute ERPs. *Lower*

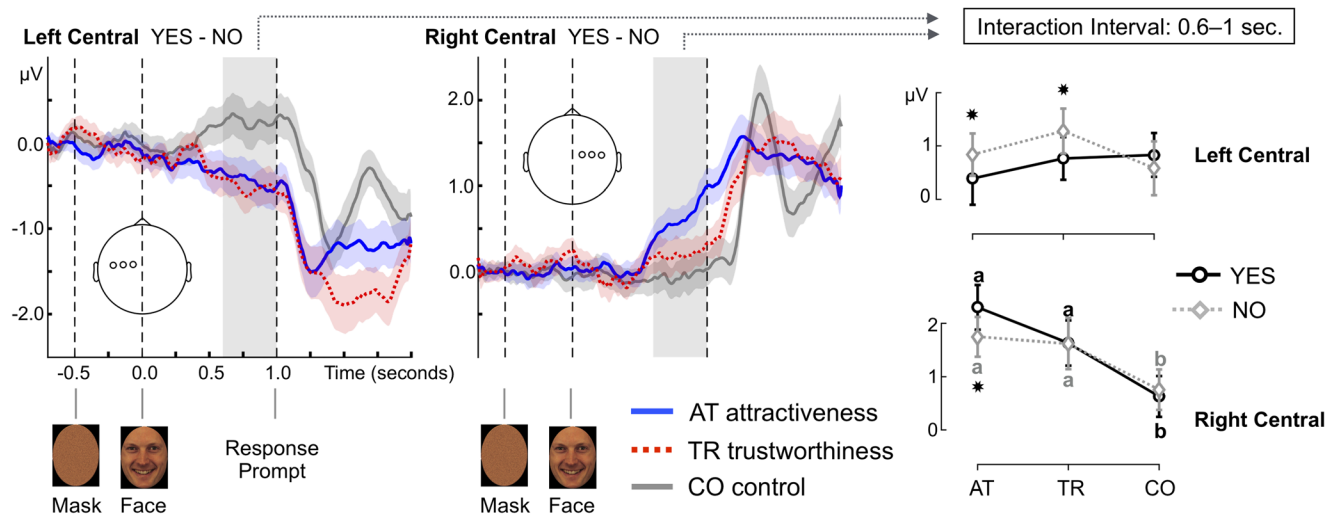
panel: ERPs elicited at centro-parietal sites as a function of *task*. The label indicates the component (P3b) that showed differences across task conditions. Colored maps represent the distribution of task differences in P3b

5.10, $p = 0.027$, $\eta_p^2 = 0.06$, was qualified by an interaction, $F(2, 72) = 4.60$, $p = 0.013$, $\eta_p^2 = 0.11$. Pairwise comparisons between Yes and No responses (to determine discrimination) revealed differences for *attractiveness* ($M_{Yes} = 0.01$ and $M_{No} = 0.56$; $SEs = 0.28$ and 0.22 ; CI s $[-0.39, 0.57]$ and $[0.17, 0.94]$), $t(24) = 2.34$, $p = 0.028$, Cohen's $d = 0.46$, CI $[0.06, 0.88]$, and for *trustworthiness* ($M_{Yes} = 0.47$ and $M_{No} = 1.02$; $SEs = 0.24$ and 0.25 ; CI s $[0.06, 0.88]$ and $[0.58, 1.44]$), $t(24) = 3.08$, $p = 0.005$, $d = 0.62$, CI $[0.17, 0.89]$, but not for the *control* task ($M_{Yes} = 0.51$ and $M_{No} = 0.39$; $SEs = 0.25$ and 0.23 ; CI s $[0.13, 0.97]$ $[-0.19, 0.83]$). The one-way (Task) ANOVAs on the

amplitude values yielded non-significant differences across tasks for Yes ($F < 1$) and No ($F = 1.76$, $p = 0.18$) responses.

For the *right central* region (C2, C4, C6; see Fig. 3A, middle plot, and the right panel), the main effect of task, $F(2, 72) = 8.45$, $p < 0.001$, $\eta_p^2 = 0.19$, was qualified by a task by response interaction, $F(2, 72) = 6.47$, $p = 0.003$, $\eta_p^2 = 0.152$. Pairwise comparisons yielded significant differences between Yes and No responses in the *attractiveness* task ($M_{Yes} = 2.05$ and $M_{No} = 1.49$; $SEs = 0.25$ and 0.22 ; CI s $[1.43, 2.67]$ and $[1.13, 1.87]$), $t(24) = 4.84$, $p < 0.001$, $d = 0.97$, CI $[0.32, 0.79]$, but *not* in the *trustworthiness* ($M_{Yes} =$

a Difference waveforms at significant clusters from AMPLITUDE analyses



b Onset LATENCIES for s-LRP, r-LRP and difference waveforms in Right Central cluster

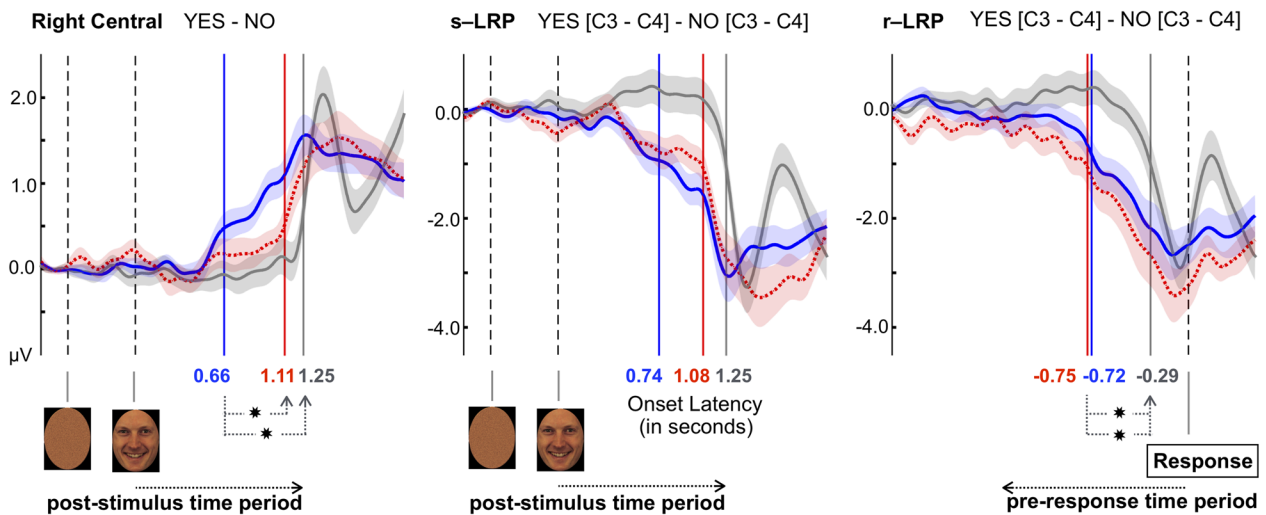


Fig. 3 Neural discrimination between *Yes* and *No* decisions. **a** *Left plot* (C1, C3, C5 sites) and *Middle plot* (C2, C4, C6 sites): Differences between ERPs elicited by faces associated with *No* vs. *Yes* responses, for attractiveness and trustworthiness tasks; and with left vs. right responses, for the control task. Grey-shaded (vertical) areas indicate the time period with a Task by Response interaction. Shaded areas of waveforms represent confidence intervals (95%). *Right panel* represents mean ERP values and CIs for *Yes* and *No* responses for the period with significant interactions (from 600 to 1,000 ms from face onset). Asterisks indicate significant differences between *Yes* and *No* responses, and letters

(a, b) reflect significant differences across tasks for each response type, *Yes* and *No*. **b** Representation of the results of the onset latency analysis for the difference waveform at right central cluster (C2, C4, C6), the stimulus-locked LRP, and the response-locked LRP. Colored vertical lines, and the numbers below, represent the mean onset latency for each task condition. Shaded areas show confidence intervals (95%). Asterisks indicate differences in latency between the tasks. The waveforms are low-pass filtered at 10 Hz, which was required to reduce high-frequency noise for the onset latency estimation

1.39 and $M_{No} = 1.38$; $SEs = 0.25$ and 0.28 ; CIs [0.95, 1.80] and [0.90, 1.85]) and the *control* ($M_{Yes} = 0.38$ and $M_{No} = 0.50$; $SEs = 0.23$ and 0.23 ; CIs [-0.01, 0.76] and [0.13, 0.89]) tasks. The one-way (Task) ANOVA, followed by post hoc contrasts (with Bonferroni corrections), showed effects for *Yes* responses, $F(2, 72) = 11.87$, $p < 0.001$, $\eta_p^2 = 0.25$, with larger

positive amplitudes for *attractiveness* ($p < 0.001$) and *trustworthiness* ($p = 0.015$) than for the *control* task. A similar pattern appeared for *No* responses, $F(2, 72) = 4.79$, $p = 0.011$, $\eta_p^2 = 0.12$, with larger positive amplitudes for *attractiveness* ($p = 0.017$) and *trustworthiness* ($p = 0.047$) than for the *control* task.

Onset latency analysis

To better characterize the discrimination time course, difference waveforms for Yes vs. No responses were computed for each task, and onset latency was estimated using the jackknife method described in the Scalp ERP analysis section. One-way (Task) ANOVAs revealed an effect for the right central region, $F(2, 72) = 12.08$, $p < 0.001$, $\eta_p^2 = 0.25$, but not for the left central region ($F = 1.51$, $p = 0.23$; Fig. 3B, left plot). At the right region, latency differences between Yes and No responses appeared earlier ($ps < 0.001$) for *attractiveness* (660 ms after face onset; $SE = 150$; CIs [484, 837]) than for *trustworthiness* (1,107 ms; $SE = 32$; CIs [930, 1284]) and the *control* task (1,251 ms; $SE = 13$; CIs [1074, 1427]), which did not differ from each other.

The previous analyses showed differences in amplitude and latency at electrodes (around C3 and C4) that are often used to examine neural activity related to response preparation and execution. To clarify the contribution of these processes to the time course differences between *attractiveness* and *trustworthiness*, the onset latency of the *stimulus-locked* LRP (Fig. 3B, middle plot) and *response-locked* LRP (Fig. 3B, right plot) components was submitted to separate one-way (Task) ANOVAs (see Scalp ERP analysis section). These analyses yielded an effect for the *response-locked* LRP, $F(2, 72) = 10.59$, $p < 0.001$, $\eta_p^2 = 0.23$, but not for the *stimulus-locked* LRP, $F(2, 72) = 2.21$, $p = 0.058$. As expected, the onset of the response-locked LRP occurred earlier for both *attractiveness* (-718 ms before response onset; $SE = 35$; CIs [-887, -560]) and *trustworthiness* (-751 ms; $SE = 126$; CIs [-909, -592]) than for the *control* condition (-288 ms; $SE = 44$; CIs [-446, -130]), $ps < 0.001$. In sum, the *attractiveness* and the *trustworthiness* tasks yielded comparable stimulus-locked LRP activity and, most importantly, their response-locked LRP activity was equivalent. Therefore, response preparation processes were not responsible for the observed time-course amplitude and onset latency differences between attractiveness and trustworthiness judgments.

Brain source localizations

To identify the likely neural contributors to scalp differences between attractiveness and trustworthiness, brain source analyses were conducted (see *Brain source analysis*, above). Neural sources were estimated for each participant and response type (Yes/No) using as input the average scalp activity from all the EEG (60) electrodes for the critical period between 600 and 1,000 ms. Next, *t*-test source maps were computed by comparing Yes and No responses for each task (t -value ≥ 3). Figure 4 (upper panels) depicts the *t*-test maps for attractiveness and trustworthiness judgments.

For *attractiveness*, several brain sources discriminated between Yes and No responses, involving regions of the *facial*

expression processing network in the right hemisphere: Lingual Gyrus (right LG; maximal in $x = 23$, $y = -69$, $z = -5$, Talairach Coordinates, corresponding to BA19; Talairach & Tournoux, 1988), Fusiform Gyrus (right FG: $x = 50$, $y = -43$, $z = -12$, BA37), and Inferior Parietal Lobe (right IPL: $x = 43$, $y = -41$, $z = 30$, BA40). Also, regions associated with *action selection and decision* processes were involved: Left Premotor Cortex (left PREM: $x = -23$, $y = 25$, $z = 51$, BA8), right Premotor Cortex (right PREM: $x = -30$, $y = -6$, $z = 42$, BA6), and **Supplementary Motor Area** (SMA, $x = -3$, $y = -7$, $z = 53$, BA6). For these brain sources, mixed 3 (Task) \times 2 (Response) ANOVAs of maximal current density values yielded significant interactions, all $F_s(2, 72) \geq 6.74$, $ps < 0.005$, $\eta_p^2 \geq 0.16$, which reflected stronger activations for Yes than for No responses in attractiveness judgments, all $t_s(24) \geq 4.23$, $ps < 0.001$, $d_s \geq 0.85$, CI for the smallest difference [0.70, 2.05], but not in the other two tasks (lower panels of Fig. 4).

For *trustworthiness*, the *t*-test map revealed an above-threshold neural source in the right Inferior Frontal Gyrus (right IFG, $x = 57$, $y = 11$, $z = 15$, BA44). In this region, the Task \times Response interaction reached significance, $F(2, 72) = 3.91$, $p = 0.025$, $\eta_p^2 = 0.10$, with stronger activations for Yes than for No responses in the trustworthiness task, $t(24) = 3.95$, $p < 0.001$, $d = 0.79$, CI [0.38, 1.21], but not in the attractiveness task. The *control* task did not produce any above-threshold (t value ≥ 3) differences.

Discussion

We investigated the neural underpinnings for the observed relationship between perception of facial attractiveness and that of trustworthiness. Our major goal was to examine the neural time course of *attractiveness* relative to *trustworthiness* processing. This serves to explore whether attractiveness impressions precede—and therefore may prime and influence—trustworthiness judgments. Complementary aims were concerned with (a) the neural time course of attractiveness and trustworthiness processing relative to emotional *facial expression* processing, and the possible interactions between attractiveness or trustworthiness and expression; (b) whether a neural discrimination advantage for un/attractiveness (or un/trustworthiness) reflects truly *encoding* processes of the face stimulus significance rather than mere response preparation processes; and (c) the *brain sources* underlying the processing time course advantage for attractiveness or trustworthiness.

Nature of the neural time course advantage

The neural processing of attractiveness and trustworthiness can be considered within the context of the widely

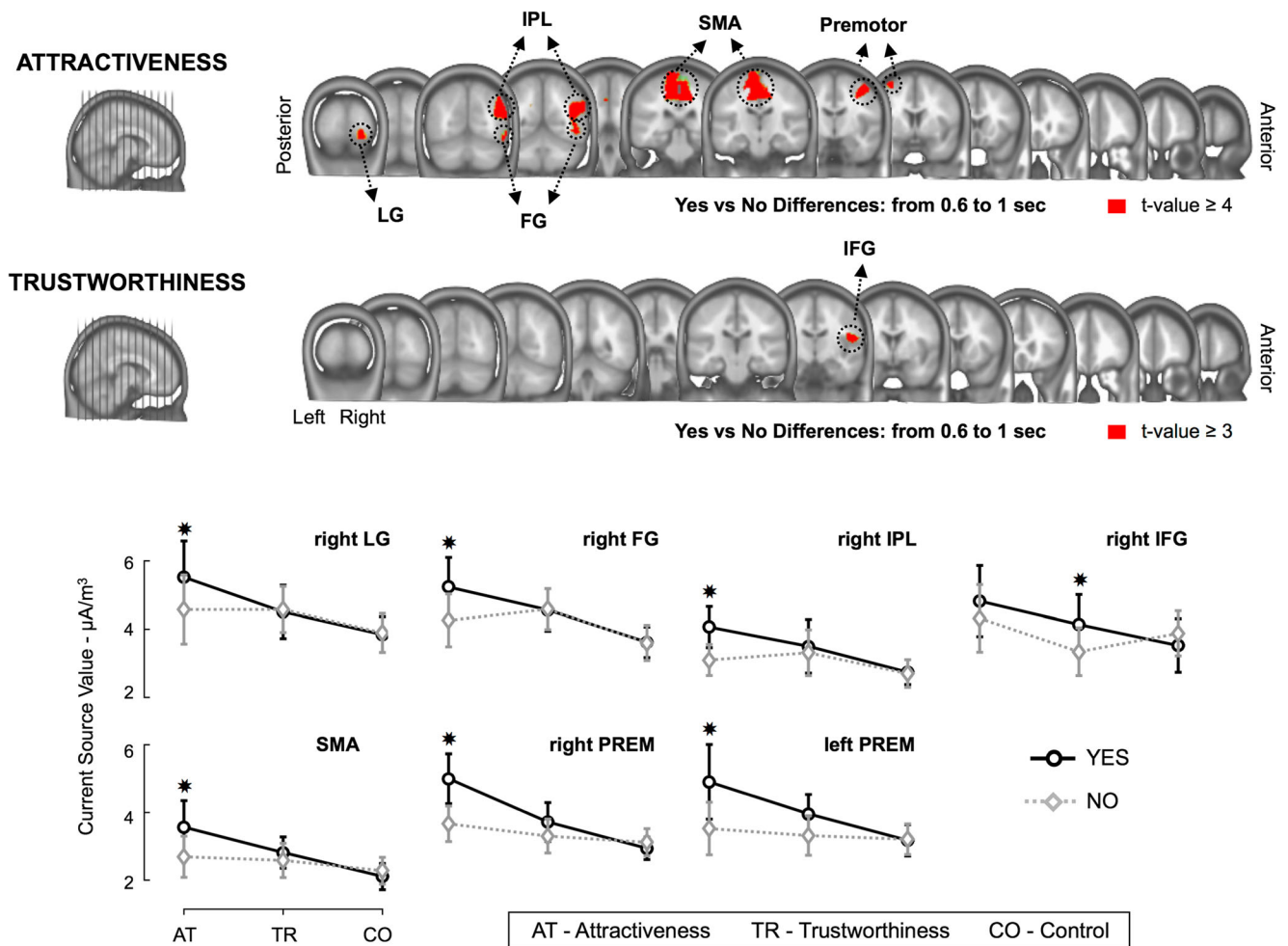


Fig. 4 Brain sources of scalp ERP activity. *Upper panel.* Brain sources for contrasts between Yes and No responses in attractiveness and trustworthiness tasks, between 600 and 1,000 ms following face onset. Red color in anatomical slices indicates regions showing significant t -values. Labels refer to neural sources: LG (Lingual Gyrus), FG

(Fusiform Gyrus), IPL (Inferior Parietal Lobe), SMA (Supplementary Motor Area), PRE (Premotor Cortex), and IFG (Inferior Frontal Gyrus). *Lower panel.* Plots represent mean current density magnitudes (and CIs) for Yes and No responses at significant ROIs. Asterisks indicate significant differences between Yes and No responses

investigated time course of emotional facial expression processing in prior research. In the current study, *expression* modulated N170 and EPN at early stages (150 to 290 ms post-stimulus onset): Faces conveying happiness, anger, and surprise enhanced neural activity in comparison with neutral faces. This is consistent with research indicating that N170 can be sensitive to emotional expression (see a meta-analysis by Hinojosa, Mercado, & Carretié, 2015) and that emotional expression typically modulates EPN (Calvo & Beltrán, 2013; Schupp et al., 2004). Furthermore, these findings support the hypothesis that emotional processing is automatic (actually, expression was task-irrelevant in the current study, yet it modulated brain activity) and occurs early and regardless of task goals (Rellecke, Sommer, & Schacht, 2012). In contrast with these early effects of expression, those of *task* (i.e., attractiveness and trustworthiness judgments vs. control condition), appeared *later*, as shown by P3b (400–700 ms; Beltrán & Calvo, 2015). Neural activity on P3b was enhanced by attractiveness

and trustworthiness relative to the control condition, and this occurred regardless of facial expression.

Furthermore, partially overlapping in time with the P3b modulation, right centro-parietal (C4, C5, C6 sites) positivity revealed that *attractiveness* was processed *faster* than trustworthiness. This was shown by the comparison of the neural activity preceding Yes vs. No response decisions (i.e., “attractive” vs. “unattractive”; or “trustworthy” vs. “untrustworthy”): Both the amplitude and latency analyses of right central activity revealed that, from 600 ms (post-stimulus onset) onwards, positivity was higher for Yes than for No decisions in the attractiveness task, relative to the trustworthiness task. Neural differences between Yes and No decisions for a given judgment (e.g., un/attractiveness) would indicate that sufficient information has been accumulated for discrimination. These findings thus demonstrated a temporal precedence for attractiveness processing. Nevertheless, effects occurred at a relatively late stage, and just preceding a

response stage, which raises the issue of whether they reflect truly stimulus encoding or, rather, motor preparation. Against this second alternative, latency-based LRP activity was equivalent for attractiveness and trustworthiness during the critical period from 600 to 1,000 ms. Such an attractiveness/trustworthiness equivalence occurred not only for stimulus-locked LRP activity, but, most importantly, for response-locked LRP activity. As LRP indexes response selection and preparation (Smulders & Miller, 2012), the LRP equivalence rules out the response preparation hypothesis to account for the attractiveness discrimination time-course advantage.

To identify the neural processes contributing to scalp differences between attractiveness and trustworthiness, we explored the underlying *brain sources* during the 600–1,000-ms period. Results support the hypothesis that the attractiveness advantage involves *face processing*: (a) the neural sources (FG, IPL—extending to pSTS—and LG) that contributed to the attractiveness scalp advantage typically underpin face or facial expression encoding (see meta-analyses by Fusar-Poli et al., 2009; Sabatinelli et al., 2011; and Vytal & Hamann, 2010); and (b) all those sources were lateralized towards the *right* hemisphere (whereas the response preparation sources—SMA and PREM—were bilateral or medial). Importantly, this shows *lateralization consistency*, as the scalp activity differences between attractiveness and trustworthiness appeared also at *right* centro-parietal sites. This right-lateralized neural activity in configural facial expression processing is also consistent with prior research using different techniques: fMRI (e.g., Maurer et al., 2007), EEG (e.g., Calvo & Beltrán, 2014), and TMS (e.g., Renzi et al., 2013). Interestingly, the neural contributor to trustworthiness in the current study came also from another right-lateralized expression processing source (IFG; see Beltrán & Calvo, 2015; Uono et al., 2017). Understandably, sources related to response preparation also were engaged (SMA and PREM), given the temporal proximity of the response stage. The important point is, however, that activity in the face processing sources was strong enough as to emerge above those involving response preparation. The greater engagement of SMA and premotor networks in the attractiveness than in the trustworthiness task presumably reveals that *decisions* about *what to respond* were ready earlier, precisely because attractiveness *judgments* occurred earlier.

Relationships with prior research

Some prior ERP studies have investigated *when* facial *trustworthiness* is processed: (a) very early C1 (40–90 ms post-stimulus onset: Yang et al., 2011), N1 (70–130 ms; Ohmann et al., 2016), or P1 (110–130 ms: Marzi et al., 2014; or P1 onset, 65–90 ms: Shore et al., 2017) modulations by un/trustworthiness have been reported, as well as (b) early N170 (140–200 ms: Dzelhyova et al., 2012) and EFP (130–220

ms; Marzi et al., 2014; 130–170 ms: Shore et al., 2017) modulations, followed by (c) later N2 (180–300 ms; Ohmann et al., 2016), EPN (230–280 ms: Dzelhyova et al., 2012; 220–400 ms: Marzi et al., 2014), and LPP/LPC (300–500 ms: Marzi et al., 2014; 400–600 ms: Yang et al., 2011; 500–800 ms: Lischke et al., 2018) effects. Our own results are in line with those showing a relatively late modulation of LPP/LPC, as shown for P3b (400–700 ms) and the discrimination between Yes (“trustworthy”) and No (“untrustworthy”) decisions from 600 ms onwards, relative to the control condition. To explain the discrepancies, it must be noted that, in studies showing early effects (Dzelhyova et al., 2012; Marzi et al., 2014; Ohmann et al., 2016; Shore et al., 2017; Yang et al., 2011), faces were generated (Marzi et al., 2014; Ohmann et al., 2016; Yang et al., 2011) or transformed (Dzelhyova et al., 2012), to enhance un/trustworthiness appearance; or, at least, they were preselected as high- or low-trust faces in a prior rating study (Shore et al., 2017). In contrast, we used natural, untransformed faces freely varying in trustworthiness. Enhanced un/trustworthiness in some prior studies could have thus speeded up discrimination (albeit at the cost of reducing naturalness and generalizability).

Numerous ERP studies have investigated *when* facial *attractiveness* is processed. In some of them, attractiveness modulated neural components earlier than 200 ms (such as P1, N1, N170, VPP; Ma & Hu, 2015; Marzi & Viggiano, 2010; Ohmann et al., 2016; Trujillo et al., 2014). Nevertheless, the early ERP effects (particularly, C1, P1, and N1, for attractiveness, and also for trustworthiness) are sensitive to low-level stimulus properties (Bieniek, Frei, & Rousselet, 2013), which were not strictly controlled in these studies. Thus, the possibility that such ERP effects might be due to physical factors, or distinctiveness, rather than attractiveness per se, cannot be ruled out (Wiese et al., 2014). Also, despite the early effects, the neural processing of attractiveness may not be mandatory, as they appeared in an explicit attractiveness judgment task, but not in a gender classification task (Schacht et al., 2008). More consistent effects, however, have been observed at later stages on N2 (~180–300 ms; Ma & Hu, 2015; Ohmann et al., 2016), EPN (~250–400 ms; Marzi & Viggiano, 2010; Thiruchselvam et al., 2016; Werheid et al., 2007; Wiese et al., 2014), and/or LPP (~400–700 ms; in practically all the studies; which is consistent with our P3b findings). Nonetheless, prior research has investigated the temporal dynamics of attractiveness and trustworthiness processing separately. Our study *compared* attractiveness and trustworthiness tasks within the same paradigm and for the same stimuli to estimate their *relative* time course. With this approach, we have shown an earlier neural processing of attractiveness, which converges with behavioral findings showing lower detection thresholds and shorter decision latencies for attractiveness than for trustworthiness judgments (Gutiérrez-García, Beltrán, & Calvo, 2018).

Theoretical implications and limitations

The neural time course advantage of *attractiveness* relative to trustworthiness suggests that an earlier first impression or evaluation of a person as un/attractive may predispose observers to judge that person as un/trustworthy, which could account for the typical facial attractiveness/trustworthiness correlation (Gutiérrez-García et al., 2018; Oosterhof & Todorov, 2008; Xu et al., 2012). There is a cognitive-economy explanation for the earlier processing of attractiveness: Reliance on easily observable attractiveness cues would serve as a mental shortcut to trustworthiness, and thus it would ease the cognitive load of making complex trustworthiness decisions about less accessible information. Relatedly, the neural time course advantage of *emotional expression* processing, relative to attractiveness/trustworthiness, implies that initially perceiving a face as angry or happy could bias viewers to judge the expresser as unattractive or attractive (and then as un/trustworthy). This is based on the fact that some expressions are consistently associated with attractiveness/trustworthiness (happy) or unattractiveness/untrustworthiness (angry) (Gutiérrez-García et al., 2018; Said et al., 2011; Sutherland et al., 2017). Thus, emotional facial expressions would modulate our first impressions for beauty and trust.

It often is assumed that human observers *automatically* infer trustworthiness from facial appearance (Bonafon, Hopfensitz, & De Neys, 2013). First, as indicated above, ERP modulations by face trustworthiness have sometimes been found at ~100 ms post-stimulus onset (Marzi et al., 2014; Ohmann et al., 2016; Yang et al., 2011). Second, detection of facial trustworthiness can occur after very short picture exposure (100 ms), backwardly masked (De Neys, Hopfensitz, & Bonafon, 2017), which suggests that trustworthiness detection is an intuitive process not requiring deliberate reflection (Bonafon, Hopfensitz, & De Neys, 2017). Third, fMRI (functional magnetic resonance imaging) research indicates that the amygdala is responsive to facial trustworthiness in the absence of perceptual awareness (Freeman, Stolier, Ingbreetsen, & Hehman, 2014). In contrast, our data do not support a strictly automatic view, as neural discrimination between trustworthy and untrustworthy faces took time to develop (from 600 to 1,000 ms). This cannot be attributed to a lack of sensitivity of the current paradigm, because it was suitable to detect the neural modulation of early ERP components (N170 and EPN) by emotional expression across all three task conditions. A difference between the current study and those of De Neys et al. (2017), Freeman et al. (2014), Marzi et al. (2014), Ohmann et al. (2016), and Yang et al. (2011) is, nevertheless, that they used face stimuli selected or manipulated to enhance un/trustworthiness, which probably facilitated discrimination.

Instead, we opted for real face stimuli and natural expressions to reflect the variability we see in everyday life. But this entailed some potential *limitations*. First, we manipulated emotional expression, but not attractiveness and trustworthiness (although their range and variability were large). This probably produced facial signals that were stronger for emotional expression than for attractiveness/trustworthiness, hence the earlier effects of emotional expression. Second, trustworthiness and attractiveness were more associated with some expressions than with others. It is thus possible that un/attractiveness and un/trustworthiness discrimination might have been influenced by expression. However, a possible confound with emotional expression can be ruled out regarding N170, EPN, and P3b, as anger, surprise, and happiness produced *equivalent* N170 and EPN (and greater than neutral faces), and the effects of task on P3b occurred *regardless* of expression, yet expressions *differed* in trustworthiness and attractiveness, with no interaction. An ideal approach should, nevertheless, use face stimuli in which expression, attractiveness, and trustworthiness were orthogonal (if this were feasible with real faces and natural expressions). Third, temporal precedence (of attractiveness processing) does not imply a causal influence (on trustworthiness processing), although it is logically incompatible with the alternative hypothesis of a reverse influence. Importantly, we were interested in the neural activity leading to conscious and intentional judgments of attractiveness and trustworthiness. Although this probably delayed the observers' decisions (and might bypass automatic processing), the crucial point is that such a delay was longer for trustworthiness than for attractiveness.

Conclusions

Emotional facial expression modulates neural activity earlier (N170 and EPN, 150–290 ms post-stimulus onset) than both facial attractiveness and trustworthiness do (P3b, 400–700 ms). Rather than merely co-occurring at a neural level, facial attractiveness processing precedes trustworthiness processing, as shown by both onset latency data and amplitude data. Neural discrimination between “Yes” (e.g., “attractive”) and “No” (e.g., “unattractive”) decisions occurred earlier (400 ms, at least) for un/attractiveness than for un/trustworthiness in the right hemisphere, with no differences in the left hemisphere. Typical facial processing brain areas (e.g., IPL and FG), also right-lateralized, were involved in the observed neural time course differences. Importantly, LRP activity showing response preparation processes was equivalent for attractiveness and trustworthiness. This suggests that the neural time course differences reflect truly facial encoding processes. The time course advantage of attractiveness discrimination suggests that it might prime (and possibly bias) trustworthiness judgments.

Acknowledgements The authors thank Patricia Alvarez-Plaza for their help in running the experiment.

Funding This research was funded by the Spanish Ministerio de Economía y Competitividad, Grant PSI2014-54720-P, awarded to Manuel G. Calvo.

Compliance with ethical standards

Disclosure of interests The authors declare no conflict of interests.

References

- Beltrán, D., & Calvo, M. G. (2015). Brain signatures of perceiving a smile: Time course and source localization. *Human Brain Mapping, 36*, 4287–4303.
- Bieniek, M. M., Frei, L. S., & Rousselet, G. A. (2013). Early ERPs to faces: Aging, luminance, and individual differences. *Frontiers in Psychology, 4*, 268. <https://doi.org/10.3389/fpsyg.2013.00268>
- Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2013). The modular nature of trustworthiness detection. *Journal of Experimental Psychology: General, 142*, 143–150.
- Bonnefon, J. F., Hopfensitz, A., & De Neys, W. (2017). Can we detect cooperators by looking at their face? *Current Directions in Psychological Science, 26*, 276–281.
- Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience, 2011*, 1–15. 813870. <https://doi.org/10.1155/2011/813870>
- Bzdok, D., Langner, R., Caspers, S., Kurz, F., Habel, U., Zilles, K., et al. (2011). ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Structure and Function, 215*, 209–223.
- Calvo, M. G., & Beltrán, D. (2013). Recognition advantage of happy faces: Tracing the neurocognitive processes. *Neuropsychologia, 51*, 2051–2060.
- Calvo, M. G., & Beltrán, D. (2014). Brain lateralization of holistic versus analytic processing of emotional facial expressions. *NeuroImage, 92*, 237–247.
- De Neys, W., Hopfensitz, A., & Bonnefon, J. F. (2017). Split-second trustworthiness detection from faces in an economic game. *Experimental Psychology, 64*, 231–239.
- De Vega, M., Beltrán, D., García-Marco, E., & Marrero, H. (2015). Neurophysiological traces of the reader's geographical perspective associated with the deictic verbs of motion to go and to come. *Brain Research, 1597*, 108–118.
- Dzhelyova, M., Perrett, D. I., & Jentsch, I. (2012). Temporal dynamics of trustworthiness perception. *Brain Research, 1435*, 81–90.
- Faugeras, F., & Naccache, L. (2016). Dissociating temporal attention from spatial attention and motor response preparation: A high-density EEG study. *NeuroImage, 124*, 947–957.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
- Freeman, J. B., Stolier, R. M., Inbretsen, Z. A., & Hehman, E. A. (2014). Amygdala responsivity to high-level social information from unseen faces. *Journal of Cognitive Neuroscience, 34*, 10573–10581.
- Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., et al. (2009). Functional atlas of emotional faces processing: A voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry and Neuroscience, 34*, 418–432.
- Grave de Peralta, R., González, S., Lantz, G., Michel, C. M., & Landis, T. (2001). Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topography, 14*, 131–137.
- Gutiérrez-García, A., Beltrán, D., & Calvo, M. G. (2018). Facial attractiveness impressions precede trustworthiness inferences: Lower detection thresholds and faster decision latencies. *Cognition and Emotion*. <https://doi.org/10.1080/02699931.2018.1444583>
- Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In S. J. Luck & E. S. Kappenman (eds.), *The Oxford handbook of event-related potential components* (pp. 441–474). New York: Oxford University Press.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223e233.
- Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 55*, 498–509.
- Lischke, A., Junge, M., Hamm, A. O., & Weymar, M. (2018). Enhanced processing of untrustworthiness in natural faces with neutral expressions. *Emotion, 18*, 181–189.
- Lundqvist, D., Flykt, A., & Öhman, A. 1998. *The Karolinska Directed Emotional Faces – KDEF*. CD-ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, Stockholm, Sweden. ISBN 91-630-7164-9.
- Ma, Q., & Hu, Y. (2015). Beauty matters: Social preferences in a three-person ultimatum game. *PLoS ONE, 10*(5): e0125806. <https://doi.org/10.1371/journal.pone.0125806>
- Marzi, T., Righi, S., Ottonello, S., Cincotta, S. M., & Viggiano, M. P. (2014). Trust at first sight: Evidence from ERPs. *Social Cognitive and Affective Neuroscience, 9*, 63–72.
- Marzi, T., & Viggiano, M. P. (2010). When memory meets beauty: Insights from event-related potentials. *Biological Psychology, 84*, 192–205.
- Maurer, D., O'Craven, K. M., Le Grand, R., Mondloch, C. J., Springer, M. V., Lewis, T. L., Grady, C. L. (2007). Neural correlates of processing facial identity based on features versus their spacing. *Neuropsychologia, 45*, 1438–1451.
- Mende-Siedlecki, P., Said, C. P., & Todorov, A. (2013). The social evaluation of faces: A meta-analysis of functional neuroimaging studies. *Social Cognitive and Affective Neuroscience, 8*, 285–299.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & de Peralta, R. G. (2004). EEG source imaging. *Clinical Neurophysiology, 115*, 2195–2222.
- Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology, 35*, 99–115.
- Ohmann, K., Stahl, J., Mussweiler, T., & Kedia, G. (2016). Immediate relativity: EEG reveals early engagement of comparison in social information processing. *Journal of Experimental Psychology: General, 145*, 1512–1529.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). Fieldtrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 1–9. 156869. <https://doi.org/10.1155/2011/156869>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the USA, 105*, 11087–11092.
- Polich, J. (2012). Neuropsychology of P300. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 159–188). New York: Oxford University Press.
- Rellecke, J., Sommer, W., Schacht, A. (2012). Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. *Biological Psychology, 90*, 23–32.
- Renzi, C., Schiavi, S., Carbon, C. C., Vecchi, T., Silvanto, J., & Cattaneo, Z. (2013). Processing of featural and configural aspects of faces is

- lateralized in dorsolateral prefrontal cortex: A TMS study. *NeuroImage*, 74, 45–51.
- Rossion, B., & Jacques, C. (2012). N170: Understanding the time course of face perception in the human brain. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 115–142). New York: Oxford University Press.
- Rudoy, J. D., & Paller, K. A. (2009). Who can you trust? Behavioral and neural differences between perceptual and memory-based influences. *Frontiers in Human Neuroscience*, 3, 16. <https://doi.org/10.3389/neuro.09.016.2009>
- Ruz, M., Madrid, E., & Tudela, P. (2013). Interactions between perceived emotions and executive attention in an interpersonal game. *Social, Cognitive, and Affective Neuroscience*, 8, 838–844.
- Sabatini, D., Fortune, E. E., Li, O., Siddiqui, A., Krafft, C., Oliver, W. T., Beck, S., & Jeffries, J. (2011). Emotional perception: Meta-analyses of face and natural scene processing. *NeuroImage*, 54, 2524–2533.
- Said, C. P., Haxby, J. V., & Todorov, A. (2011). Brain system for assessing the affective value of faces. *Philosophical Transactions of the Royal Society*, 366, 1660–1670.
- Schacht, A., Werheid, K., & Sommer, W. (2008). The appraisal of facial beauty is rapid but not mandatory. *Cognitive, Affective and Behavioral Neuroscience*, 8, 132–142.
- Schupp, H., Öhman, A., Junghöfer, M., Weike, A. I., Stockburger, J., & Hamm, A. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion*, 4, 189–200.
- Shore, D. M., Ng, R., Bellugi, U., & Mills, D. L. (2017). Abnormalities in early visual processes are linked to hypersociability and atypical evaluation of facial trustworthiness: An ERP study with Williams syndrome. *Cognitive, Affective, and Behavioral Neuroscience*, 17, 1002–1017.
- Smulders, F. T. (2010). Simplifying jackknifing of ERPs and getting more out of it: Retrieving estimates of participants' latencies. *Psychophysiology*, 47, 387–392.
- Smulders, F. T., & Miller, J. O. (2012). The lateralized readiness potential. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 209–229). New York: Oxford University Press.
- Sutherland, C. A. M., Rowley, L., Amoaku, U., Daguzan, E., Kidd-Rossiter, K., Maceviciute, U., & Young, A. W. (2015). Personality judgements from everyday images of faces. *Frontiers in Psychology*, 6, 1616.
- Sutherland, C. A. M., Young, A. W., & Rhodes, G. (2017). Facial first impressions from another angle: How social judgements are influenced by changeable and invariant facial properties. *British Journal of Psychology*, 108, 397–415.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Thieme: Stuttgart.
- Thiruchselvam, R., Harper, J., & Homer, A. L. (2016). Beauty is in the belief of the beholder: Cognitive influences on the neural response to facial attractiveness. *Social Cognitive and Affective Neuroscience*, 11, 1999–2008.
- Todorov, A. T., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, 66, 519–545.
- Trujillo, L. T., Jankowitsch, J. M., Langlois, J. H. (2014). Beauty is in the ease of the beholding: A neurophysiological test of the averageness theory of facial attractiveness. *Cognitive, Affective, and Behavioral Neuroscience*, 14, 1061–1076.
- Uono, S., Sato, W., Kochiyama, T., Sawada, R., Kubota, Y., Yoshimura, S., et al. (2017). Neural substrates of the ability to recognize facial expressions: A voxel-based morphometry study. *Social, Cognitive, and Affective Neuroscience*, 12, 487–495.
- van Hooff, J. C., Crawford, H., & van Vugt, M. (2011). The wandering mind of men: ERP evidence for gender differences in attention bias towards attractive opposite sex faces. *Social, Cognitive, and Affective Neuroscience*, 6, 477–485.
- Vytal, K., & Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: A voxel-based meta-analysis. *Journal of Cognitive Neuroscience*, 22, 2864–2885.
- Wegrzyn, M., Riehle, M., Labudda, K., Woermann, F., Baumgartner, F., Pollmann, S., et al. (2015). Investigating the brain basis of facial expression perception using multi-voxel pattern analysis. *Cortex*, 69, 131–140.
- Werheid, K., Schacht, A., & Sommer, W. (2007). Facial attractiveness modulates early and late event-related brain potentials. *Biological Psychology*, 76, 100–108.
- Wiese, H., Altmann, C. S., & Schweinberger, S. R. (2014). Effects of attractiveness on face memory separated from distinctiveness: Evidence from event-related brain potentials. *Neuropsychologia*, 56, 26–36.
- Xu, F., Wu, D., Toriyama, R., Ma, F., Itakura, S., & Lee, K. (2012). Similarities and differences in Chinese and Caucasian adults' use of facial cues for trustworthiness judgements. *PLoS ONE* 7:e34859. <https://doi.org/10.1371/journal.pone.0034859>
- Yang, D., Qi, S., Ding, C., Song, Y. (2011). An ERP study on the time course of facial trustworthiness appraisal. *Neuroscience Letters*, 496, 147–151.