

Beauty is in the ease of the beholding: A neurophysiological test of the averageness theory of facial attractiveness

Logan T. Trujillo · Jessica M. Jankowitsch ·
Judith H. Langlois

Published online: 11 December 2013
© Psychonomic Society, Inc. 2013

Abstract Hundreds of studies have shown that people prefer attractive over unattractive faces. But what is an attractive face, and why is it preferred? Averageness theory claims that faces are perceived as being attractive when their facial configuration approximates the mathematical average facial configuration of the population. Conversely, faces that deviate from this average configuration are perceived as being unattractive. The theory predicts that both attractive and mathematically averaged faces should be processed more fluently than unattractive faces, whereas the averaged faces should be processed marginally more fluently than the attractive faces. We compared neurocognitive and behavioral responses to attractive, unattractive, and averaged human faces to test these predictions. We recorded event-related potentials (ERPs) and reaction times (RTs) from 48 adults while they discriminated between human and chimpanzee faces. The participants categorized averaged and high-attractive faces as being “human” faster than low-attractive faces. The posterior N170 (150–225 ms) face-evoked ERP component was smaller in response to high-attractive and averaged faces than to low-attractive faces. Single-trial electroencephalographic analysis indicated that this reduced ERP response arose from the engagement of fewer neural resources, and not from a change in the temporal consistency of how those resources were engaged. These findings provide novel evidence that faces are perceived as being attractive when they approximate a facial configuration

close to the population average, and they suggest that processing fluency underlies preferences for attractive faces.

Keywords Facial attractiveness · Averaging · Event-related potentials · Visual cortex

It is well established that human adults, children, and even infants judge and behave toward others differently, depending on facial attractiveness. Unattractive children and adults are judged as being less well-adjusted, less socially appealing, and less academically and interpersonally competent than attractive children and adults (Langlois et al. 2000). Infants approach and interact with attractive strangers, but withdraw from and avoid unattractive strangers (Langlois, Roggman, and Rieser-Danner 1990). Mothers of attractive infants provide their infants with more positive attention than do mothers of unattractive infants (Langlois, Ritter, Casey, and Sawin 1995). Attractive adults earn higher wages than do unattractive adults with the same level of education (Hamermesh 2011; Judge, Hurst, and Simon 2009) and are considered to be more intelligent and sociable (Griffin and Langlois 2006). Thus, given the ubiquitous role that facial attractiveness plays in social and affective interactions, it is important to understand what makes faces attractive and why attractive faces are preferred.

Theories of facial attractiveness

Two theoretical perspectives have emerged as explanations of facial attractiveness: one based on evolutionary principles and the other based on information processing and cognitive averaging. Most evolutionary theories view preferences for attractive faces as the result of an evolved, domain-specific module that identifies—for mate selection purposes—good genes, health, and reproductive fitness. Indicators of health and reproductive fitness are different in men and women, but include square jaws (men), high cheekbones (women), big eyes and lips (women),

Electronic supplementary material The online version of this article (doi:10.3758/s13415-013-0230-2) contains supplementary material, which is available to authorized users.

L. T. Trujillo · J. M. Jankowitsch · J. H. Langlois
Department of Psychology, University of Texas at Austin, Austin,
TX, USA

L. T. Trujillo (✉)
Department of Psychology, University of Texas at Austin, 108 E.
Dean Keeton, Stop A8000, Austin, TX 78712-1043, USA
e-mail: trujillo@austin.utexas.edu

and symmetry (both), among other factors (Fink and Penton-Voak 2002; Perrett et al. 1998; Rhodes 2006; Rhodes, Proffitt, Grady, and Sumich 1998; Thornhill and Gangestad 1993, 1999; but see Kalick, Zebrowitz, Langlois, and Johnson 1998). A major argument against these particular evolutionary accounts is that although health and reproductive fitness indicators may enhance attractiveness, these characteristics are neither necessary nor sufficient to produce it. Many perfectly symmetrical faces are not attractive, many youthful faces are not attractive, many faces with big lips are not attractive, and many highly masculine faces are not attractive (Langlois, Roggman, and Musselman 1994; Rubenstein, Langlois, and Roggman 2002).

In contrast, theoretical perspectives based on cognitive averaging (collectively termed here as *averageness theory*) view preferences for attractive faces as resulting from their similarity to *facial prototypes*—the categorical central tendencies of a population of faces (Langlois and Roggman 1990). In this account, facial attractiveness is driven by a face's *averageness*, or the degree to which its configuration approximates the mathematical average facial configuration of a population of male or female faces. Multiple studies have shown that adult humans perceive faces possessing high averageness as being highly attractive, whereas faces with low averageness (i.e., faces that deviate from the average facial configuration of a population) are perceived as being unattractive (Bronstad, Langlois, and Russell 2008; Langlois and Roggman 1990; Langlois, Roggman, Casey, and Ritter 1987; Langlois et al. 1994; Rhodes and Tremewan 1996; Rubenstein, Kalakanis, and Langlois 1999). This preference for average faces is shared across cultures (Apicella, Little, and Marlowe 2007). These studies suggest that the averageness of a face is both necessary *and* sufficient for determining facial attractiveness (Langlois et al. 1994; Rubenstein et al. 2002). Preferences for faces near their prototype arise because, in general, prototypical stimuli are processed more rapidly and efficiently than other stimuli (Komatsu 1992; Posner and Keele 1970; Reed 1972; Rosch, Mervis, Gray, Johnson, and Boyes-Braem 1976). Such fluent stimulus processing is associated with positive affective responses that render prototypical stimuli more visually “pleasing” than nonprototypical stimuli (Harmon-Jones and Allen 2001; Principe and Langlois 2011, 2012; Winkielman and Cacioppo 2001; Winkielman, Halberstadt, Fazendeiro, and Catty 2006). This is presumably because fluent processing is typically error-free and indicates successful recognition of a stimulus (Winkielman et al. 2006).

The present study

The purpose of the present study was to examine the predictions of averageness theory for the neural processing of facial attractiveness, since, to our knowledge, such an examination has not yet been performed. Our objective was not to devise a single experiment that could decide between averageness and

evolutionary theory, but instead to provide evidence to either support or falsify the neural processing predictions of the former theory. To accomplish this, we compared neurocognitive responses to averaged faces (prototypes) and to individual faces varying in attractiveness. We recorded event-related potential (ERP) measures of stimulus-evoked electroencephalographic (EEG) activity at the scalp while adults categorized images of faces as either human or chimpanzee, a task that engaged automatic, stimulus-driven processing related to the attractiveness of the human faces. Three kinds of human faces were presented: high-attractive faces, low-attractive faces, and averaged faces consisting of mathematical composites of 32 individual faces varying along a wide range of attractiveness. The averaged faces were used as estimates of the mean of the face population from which they were sampled (see the [Stimuli and Procedure](#) section, below) and are highly attractive, not average, in appearance (see the [Behavior](#) section in the Results, below, and the [supplementary materials](#)).

We tested two predictions of averageness theory in this study. First, attractive faces should be perceived as more typical and less distinctive than unattractive faces, since the former are more similar to the prototype, whereas the latter are more distinguishable from the prototype. This prediction had been supported by studies that revealed attractiveness to be negatively correlated with distinctiveness (Light, Hollander, and Kayra-Stuart 1981; Peskin and Newell 2004; Rhodes and Tremewan 1996; Vokey and Read 1992). Here we expected that both averaged and attractive faces would be rated as less distinctive than low-attractive faces. However, we also expected that averaged faces would be rated as less distinctive than high-attractive faces, because the former are closer to the face prototype than are the latter. Second, attractive faces should be processed more fluently than unattractive faces, because of their similarity to the facial prototype. Consistent with this prediction, Hoss, Ramsey, Griffin, and Langlois (2005) found in a reaction time (RT) study that attractiveness facilitates the speed and accuracy of gender-based face classification (but see Schacht, Werheid, and Sommer 2008). In the present study, we hypothesized that participants would categorize high-attractive and averaged faces as being “human” faster than low-attractive faces, and would categorize the averaged faces marginally faster than the high-attractive faces.

We also measured processing fluency via the well-known ERP index of face processing, the N170 component.¹ The N170

¹ We also measured the early (~100 ms) P1 and the intermediate stage (~200–300 ms) P2 ERP components as part of our N170 quantification methodology (see the [ERP Quantification](#) section in the Method and the [N170 Latency Jitter Correction Analysis](#) section in the Results, below). Despite a previous finding that the P1 component is sensitive to facial attractiveness (Halit, de Haan, and Johnson 2000), we do not present a detailed analysis of the P1 component in this article, because preliminary analyses revealed no P1 differences among the human faces. In the supplemental materials, we report an analysis of the P2, since this component reflected attractiveness-related affective factors that were not directly relevant to processing fluency, the main focus of this article.

is a negative-polarity ERP component observed over bilateral occipital temporal scalp regions approximately 170 ms post-stimulus-onset. The N170 response is larger for faces than for nonface objects (Bentin, Allison, Puce, Perez, and McCarthy 1996; Rossion and Jacques 2011), although in general this component is sensitive to entry-level object categorization, of which face perception is a privileged case (Rossion, Curran, and Gauthier 2002). The N170 has a magnetoencephalographic counterpart, the M170, which has a similar time course and response properties (Deffke et al. 2007; Halgren, Raji, Marinkovic, Jousmäki, and Hari 2000; Linkenkaer-Hansen et al. 1998). Although the N170/M170 complex arises from activity in multiple cortical regions (including occipitotemporal sulcus, posterior superior temporal sulcus, posterior lingual gyrus, and the fusiform gyrus), it appears that the dominant sources are in the lateral posterior fusiform gyrus and anterior/middle fusiform gyrus (for a full review, see Rossion and Jacques 2011). The latter regions respectively correspond to the occipital and fusiform face areas, as identified by functional magnetic resonance imaging (fMRI) studies (Haxby, Hoffman, and Gobbini 2000; Kanwisher, McDermott, and Chun 1997).

Importantly, the N170 is increased and often delayed in response to inverted versus upright photorealistic faces (Bentin et al. 1996; Itier and Taylor 2002, 2004a, 2004b; Jacques and Rossion 2007; Rossion et al. 1999; Rossion and Jacques 2011; Rossion, Joyce, Cottrell, and Tarr 2003; Vizioli, Foreman, Rousselet, and Caldara 2010). These N170 face inversion effects have been attributed to the disruption of configural face processing (Rossion et al. 1999; Rossion and Gauthier 2002) because face recognition is disproportionately affected by inversion, as compared to other objects (Yin 1969). Thus the N170 ERP response provides a sensitive metric to investigate fluent versus nonfluent face processing. With respect to attractiveness, Halit, de Haan, and Johnson (2000) observed that atypical unattractive faces elicit a larger N170 ERP response than do typical attractive faces (when typicality varies across, but not within, facial identities). Halit et al. suggested that their N170 effects “may reflect processing related to perceiving an individual in relation to a general face prototype” (p. 1874), consistent with the interpretation that the N170 responses observed in their study arose from the more fluent processing of the attractive than of the unattractive faces. Nonetheless, their finding cannot be taken as a confirmation of cognitive averaging theory because they did not include averaged faces in their stimulus set or obtain behavioral measures that could index fluency of processing—hence, one motivation for the present study. On the basis of the predictions of averageness theory and previous observations of the N170, we hypothesized that low-attractive faces would elicit larger N170 amplitudes and/or more delayed latencies than high-attractive and averaged faces, whereas high-attractive faces would elicit larger N170 amplitudes and/or more delayed latencies than averaged faces.

The above-reviewed N170 findings are consistent with other evidence that fluent stimulus processing is associated with a reduction of neural activity in the brain. For example, face-selective neurons in monkey inferotemporal cortex exhibit monotonic decreases in firing rates as face stimuli transition from nonprototypical individual faces toward the average of the faces (Leopold, Bondar, and Giese 2006). Hemodynamic activity in human fusiform gyrus and occipital cortex is smaller for prototypical than for nonprototypical faces and visual dot patterns, respectively (Loffler, Yourganov, Wilkinson, & Wilson, 2005; Reber, Stark, and Squire 1998). Nonetheless, it is not yet clear *why* prototypical stimuli are processed fluently. Prototypical stimuli may require fewer neural resources to process, and thus recruit a lesser number of neurons and/or activate neurons to a lesser degree, than do nonprototypical stimuli. Then again, prototypical stimuli may engage neural resources with a rapid and consistent time course, whereas other stimuli engage the same neural responses more slowly and/or with a more variable time course. A third possibility is that processing fluency may affect both the degree and timing of neural response. Hence, distinguishing among these alternatives was an additional goal of the present study. We utilized a simple single-trial peak detection procedure (Spencer 2005; Spencer, Abad, and Donchin 2000) that allowed for separate measurement of amplitude and timing changes across conditions (see the [ERP Quantification](#) section below).

Method

Participants

A group of 55 undergraduates participated for course credit or payment. We excluded seven participants due to excessive EEG artifacts or technical recording problems (see the EEG acquisition, preprocessing, and data reduction section below). Thus, the final sample consisted of 48 undergraduates (30 female, 18 male; 18.60 ± 0.17 years of age, 44 right-handed); this number of participants was chosen a priori in order to provide balanced assignment of stimuli to each condition, as was necessitated by our experimental design (see the [Stimuli and Procedure](#) section below). All participants were fully informed of the experimental methods and proceedings before consent. The Institutional Review Board for Human Studies at the University of Texas at Austin approved this study.

Stimuli and procedure

Participants viewed color images of high-attractive, low-attractive, and averaged young adult Caucasian female faces in the context of a species categorization task, described below. (We included only female faces in this study because adult observers of both sexes more consistently agree on the attractiveness of female faces than on the attractiveness of

male faces; see, e.g., Hoss et al. 2005; Rennels, Bronstad, and Langlois 2008.) We constructed the averaged faces by mathematically averaging 32 individual faces (Langlois and Roggman 1990; Langlois et al. 1994). A constant-size oval occluding window, applied to each photograph, excluded the background and hair. All faces were presented against a uniform white background. Importantly, the low- and high-level visual properties that might differentially affect behavioral and/or ERP responses were equated across human face categories (see the [supplementary materials](#) for a description). Another group of participants ($N = 55$; see the [supplementary materials](#) for the ratings procedure) determined the a priori attractiveness of each face. Participants rated the faces along a 7-point attractiveness scale, with high-attractive and averaged faces being rated near the high end of the scale (mean ratings of 5.15 ± 0.15 and 5.60 ± 0.12 , respectively) and the low-attractive faces being rated near the low end of the scale (mean rating = 2.06 ± 0.09).

One notable feature of averaged faces is that different averaged faces constructed from different face sets are highly perceptually similar (see Fig. 1). This poses a problem when presenting averaged faces in the context of experimental designs that present exemplars of each face/stimulus category: Perceptually similar averaged faces are likely to be perceived as multiple instances of the same averaged face, which therefore could induce differential repetition effects between averaged faces and other face categories, even though all faces were presented equally often. To avoid this confound, we implemented a design that ensured that all individual human faces and general human face categories (attractive vs. unattractive) would be repeated, and perceived to be repeated, equally often. We created three groups of faces on the basis of the a priori attractiveness ratings—one group of high-attractive individual faces (two faces), one group of low-

attractive individual faces (four faces), and one group of averaged faces (two faces). These faces were rated as the most highly attractive or unattractive in their category. Each participant repeatedly viewed one face each from the high-attractive and averaged face groups, and two faces from the low-attractive face group (four faces total). The selection of faces from the face category groups was balanced across participants and categories; each participant received 1 of 48 possible unique across-category combinations of faces. Block analyses showed that potential differential habituation effects due to the numerous repetitions of each face were negligible (see the [ERP Quantification](#) section in the [Method](#) and the [Behavior](#) section in the Results, below).

Participants also viewed a set of ten color images of chimpanzee faces that were the same size as, and with the external background, hair, and head features excluded in the same manner as in, the human faces (Fig. 1). Participants individually categorized each face as human or chimpanzee as quickly and accurately as possible. This cover task served two purposes: First, it ensured that participants attended to the faces; second, it encouraged participants to appraise the human faces in a manner that did not directly involve facial attractiveness. An important aspect of our design was that ERP responses to the three categories of human faces would be assessed under conditions in which they received the same classification response, thus ensuring that any observed ERP differences were not due to differences in conscious stimulus categorization. Hence, the comparisons of primary interest in this article involved comparisons among the human faces. Nonetheless, we also report behavioral and ERP responses to the chimpanzee faces for comparison with the responses to the human faces, as well as to ascertain that participants performed the species categorization task satisfactorily (see the [Behavior](#) section in the Results, below).

Participants viewed the randomly intermixed human and chimpanzee faces over 12 trial blocks, with 50 face presentations per block (10 high-attractive, 10 averaged, 20 low-attractive, 10 chimpanzee); see Fig. 1. We presented the faces for 250 ms, centered on a 17-in. computer LCD screen. All of the faces were displayed against a white background and subtended $\sim 7.52^\circ$ (w) \times 11.31° (h) of visual angle at a viewing distance of 75 cm; a small cross was presented at central fixation during all of the interstimulus intervals (ISIs). For each face, the participants indicated a human-versus-chimpanzee categorization by using their dominant hand to press one of two buttons on a response box placed in front of them on a table (left button = human, right button = chimpanzee). The ISIs ranged from 2,250 to 3,250 ms; participants had 1,500 ms to respond on each trial.

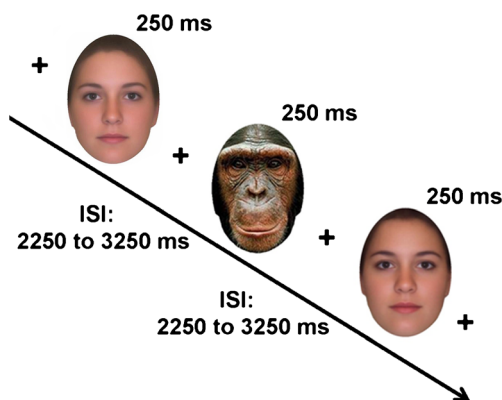


Fig. 1 Basic task design. On all trials, participants viewed a central fixation cross before a face was presented for 250 ms. The interstimulus intervals ranged from 2,250 to 3,250 ms, with the central fixation cross being displayed throughout. Participants categorized the faces, via a buttonpress, as being either human or chimpanzee. The figure shows the two averaged face exemplars used in this study; note the high perceptual similarity between the two averaged faces (see the [Stimuli and Procedure](#) section)

After the categorization task, participants rated the human faces for facial attractiveness and distinctiveness (see the [supplementary materials](#) for a description of the rating methods). Facial distinctiveness was defined as the ease of spotting a face in a crowd (Wickham and Morris 2003).

Behavioral data reduction and analysis

On the basis of the known limits of human visual processing speed and recommended methods for dealing with RT outliers (Fabre-Thorpe, Delorme, Marlot, and Thorpe 2001; Ratcliff 1993), we restricted the categorization task analysis to correct trials with RTs >200 ms and <2.5 standard deviations of the grand mean across conditions. In order to remove false starts and extreme outliers, 2.48 % ± 0.12 % of all trials were rejected due to RTs outside the specified limits. Next, the 12 experimental blocks of behavioral data were separated into four groups of three blocks for each condition (i.e., Block Group 1 = Trial Blocks 1–3, Block Group 2 = Trial Blocks 4–6, etc.). We computed the mean categorization task RTs and accuracy–hit rates (defined as the percentages of correct trials within the total number of trials, after removal of timeout trials) for each block group, face category, and participant. Separation of the data into block groups was performed to reduce data variability and simplify the data interpretation and analysis. Because twice as many low-attractive trials were presented as high-attractive and averaged trials, we computed each individual's behavioral measures for the low-attractive condition from a random sampling (without replacement) of their low-attractive trials. This sampling was constrained such that (1) the number of sampled trials for a given block group was equivalent to the average number of correct high-attractive and averaged face trials for that group, and (2) the sampled trials contained approximately equal contributions from the two sets of low-attractive face exemplar trials. In addition, we computed mean attractiveness and distinctiveness ratings from the human face rating data (low-attractive face ratings were averaged across all of the faces in this category).

We used the SPSS statistical software package (IBM Corp., Armonk, NY, USA) to compare among responses to the human and chimpanzee faces via two-way repeated measures analyses of variance (ANOVAs) with the factors Block Group and Face Type (high-attractive, averaged, low-attractive, or chimpanzee). We also analyzed the mean attractiveness and distinctiveness ratings of the human faces via one-way repeated measures ANOVAs. To account for sphericity violations when the factors involved more than two levels, the *p* values of all ANOVAs were adjusted using the Greenhouse–Geisser correction (Greenhouse and Geisser 1959). For ease of interpretation, the reports below of all significant *F* tests subject to Greenhouse–Geisser correction include uncorrected degrees of freedom, corrected *p* values, and the Greenhouse–Geisser epsilon value ϵ . In addition, the *p* values for all reported post-hoc

comparisons and auxiliary ANOVAs used to decompose significant interactions were corrected for multiple comparisons using the Holm–Bonferroni procedure (Holm 1979), which controls the family-wise error rate at level α (here, $\alpha = .05$). Effect sizes were estimated via the partial eta-squared (η^2_p) measure.

We also examined the predictive relationship between the distinctiveness and attractiveness ratings of the human faces via generalized estimating equations using SPSS (Gardiner, Luo, and Roman 2009; Ghisletta and Spini 2004; normal distribution with identity link, robust covariance estimate, exchangeable working correlation matrix, and maximum-likelihood-estimate scale parameter), a generalized regression procedure that accounts for correlations across repeated measures.

EEG acquisition, preprocessing, and data reduction

Participants wore a 32-channel Geodesics Sensor Net (Electrical Geodesics Inc. [EGI], Eugene, OR, USA), which included channels placed at the inferior orbits and outer canthi of each eye. The impedances were adjusted to below 50 k Ω , following standard recommendations for the EGI system. We recorded continuous EEG (250-Hz sampling rate) via the NetStation 4.4 software (Electrical Geodesics Inc.) with respect to a CZ reference. We extracted 2,000-ms EEG data epochs, from –750 to 1,250 ms with respect to the onset of the human face stimuli, that were imported into the MATLAB computing environment (The MathWorks, Natick, MA, USA).

We performed all subsequent analyses in MATLAB via in-house scripts that utilized functions from the EEGLAB toolbox (Delorme and Makeig 2004). First, the epoched EEG data were transformed to an average reference.² Next, we excluded incorrect trials and trials with RTs <200 ms, or >2.5 standard deviations from the grand mean across conditions, from further analysis; on average, 3.20 % ± 0.19 % and 2.03 % ± 0.28 % of the human face category and chimpanzee face category EEG trials, respectively, were rejected for these reasons. We also removed trials contaminated with muscle and EEG signal artifacts via visual inspection. Bad EEG channels were replaced using an EEGLAB-based spherical spline interpolation algorithm (Perrin, Perrier, Bertrand, Giard, and Echallier 1987; mean number of interpolated channels = 0.48 ± 0.12). We removed electrooculographic (EOG) artifacts due to blinks and saccades via a two-stage procedure (Trujillo, Kornguth,

² We utilized an average reference because Joyce and Rossion (2005) showed that it optimally indexes categorical differences for the N170 ERP component. Although the size of the electrode montage used in the present study (32 channels) was lower than is typically recommended when using an average reference (>64 electrodes; Junghöfer, Elbert, Tucker, and Braun 1999), we believe that these ERP findings are not dependent on our choice of EEG reference. An additional analysis (data not shown) utilizing a CZ reference and an estimation of a neutral reference “at infinity” (Yao 2001) revealed ERP effects similar to those found using the average reference.

and Schnyer 2009) that simultaneously minimizes data loss and the physical and neural impacts of ocular activity on the EEG signals. In the first stage, we rejected trials contaminated with EOG activity in the –200- to 300-ms interval with respect to face onset, on the basis of the activity present in two bipolar montage channels computed offline (vertical EOG: FP2–[right inferior orbit]; horizontal EOG: F9–F10; rejection criteria: $\pm 40 \mu\text{V}$). In the second stage, we corrected the scalp EEG signals for residual EOG activity outside the rejection interval by application of an adaptive-filter regression-based ocular correction algorithm (He, Wilson, and Russell 2004) to the entirety of each EEG epoch using the Automatic Artifact Removal (Version 1.3) add-on for the EEGLAB toolbox. Because ocular artifact correction algorithms occasionally fail to remove all EOG artifacts on some trials, a second round of manual artifact scoring was performed, after band-pass filtering the EEG data between 0.1 and 48 Hz (166-point zero-phase-shift FIR filter with 0.03- and 7-Hz transition bands, respectively). On average, $15.11 \% \pm 1.15 \%$ and $25.66 \% \pm 1.64 \%$ of the human face category and chimpanzee face category EEG trials, respectively, were rejected due to muscle, signal, and/or ocular artifacts.

After artifact removal, the band-pass-filtered EEG trials were truncated to the –200- to 400-ms interval, to remove artifactual filter-related edge effects. The EEG trials were then baseline-corrected to the –200- to 0-ms prestimulus interval and divided into the three human face categories (high-attractive, averaged, and low-attractive); the average number of trials per human face category was 98.02 ± 0.87 (after collapsing across all experimental blocks; see below).

ERP quantification

We visually estimated the general time window and scalp location of the N170 (150–225 ms) ERP components from the across-participant grand-average ERPs. We quantified the N170 component amplitudes at the six temporal–occipital scalp sites (P7, P8, PO9, PO10, M1, and M2) where N170 effects are typically maximal (Bentin et al. 1996; Rossion et al. 1999), as was the case in the present grand-average ERPs (see Fig. 2 in the Results). To simplify the data interpretation, the ERP waveforms at left- and right-hemisphere locations were averaged together separately to form two regional waveforms. The N170 latency was defined as the latency of the most-negative peak within the general N170 time window. The N170 amplitude was quantified as the mean activity within the time window defined by those points in the immediate leading and lagging edges of the N170 peak where the ERP amplitude was 75 % of the peak amplitude (Picton et al. 2000). We analyzed the ERP component amplitudes and latencies via two-way repeated measures univariate ANOVAs with the factors Face Type and Hemisphere, after collapsing across

experimental blocks.³ All p values were adjusted using the Greenhouse–Geisser correction for nonsphericity; post-hoc comparisons and auxiliary ANOVAs were corrected for multiple comparisons via the Holm–Bonferroni procedure; and effect sizes were estimated via the η^2_p measure (see the Behavioral Data Reduction and Analysis section, above).

For graphical display, we generated grand-average ERP waveforms by averaging the ERPs at a given electrode across participants separately for each condition. We visualized the scalp distributions of ERP component differences by computing the between-condition amplitude differences for a component at each electrode, which were displayed as interpolated topographical scalp maps using algorithms from the EEGLAB MATLAB toolbox.

N170 latency jitter correction analysis

To determine the contributions of amplitude and/or timing changes to the attractiveness-related N170 effects across the human face conditions, we identified the N170 latency on each trial and corrected the ERPs for any latency jitter that was present. By comparing N170 amplitudes after latency jitter correction, as well as the standard deviations of the single-trial peak latencies for each condition, we could ascertain whether our effects involved changes in ERP amplitude, timing, or both. To this end, we adapted a peak-picking latency jitter correction technique previously utilized to estimate P300 latencies (Spencer et al. 2000). This computationally simple technique matches the performance of more sophisticated latency estimation measures when signal-to-noise ratios are high and/or the width of the true ERP component is relatively narrow (Gratton, Kramer, Coles, and Donchin 1989; Jáskowski and Verleger 2000; Smulders, Kenemans, and Kok 1994). This method was well suited to our data, given the relatively clean ERP baselines therein (see the Event-Related Potentials section in the Results) and the high numbers of trials entering into each ERP. A limitation of our procedure is that it cannot distinguish between changes in the across-trial variability of ERP peak latencies versus changes in the rate at which ongoing prestimulus EEG oscillations are realigned following stimulation onset, a process known as *phase-resetting* (Gruber, Klimesch, Sauseng, and Doppelmayr 2005; Klimesch et al. 2004; Makeig et al. 2002; Trujillo and Allen 2007). However, this limitation did not impede our present goal of ascertaining whether EEG timing changes contribute to N170 effects of facial attractiveness by either timing mechanism.

³ Additional ANOVAs on the non-latency-jitter-corrected ERP data, which included Block as a factor, revealed no significant differences among the different human face types as a function of block for the N170 amplitudes or latencies ($p < .341$). Thus, we report the results of the N170 component derived from averages across all trials within each condition in order to increase the signal-to-noise ratio of the ERPs, and hence the statistical power.

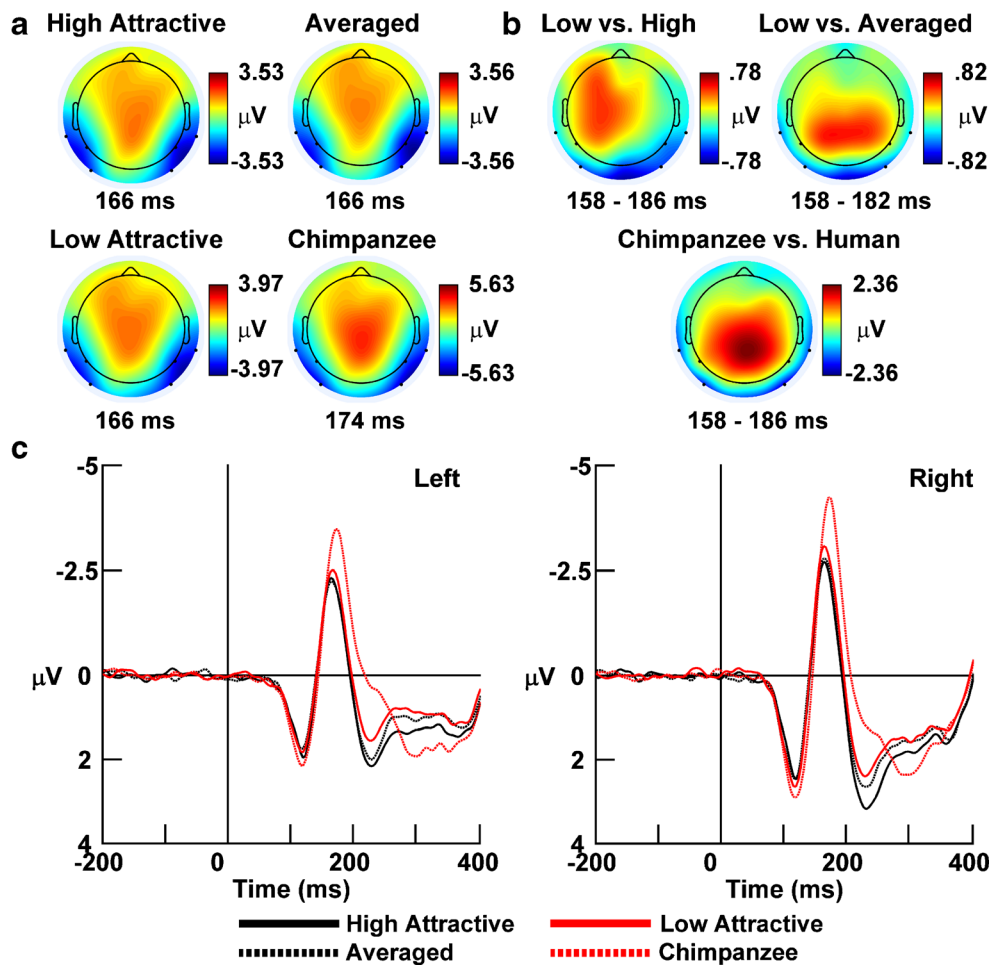


Fig. 2 (a) N170 peak and (b) N170 difference wave scalp topographies for all four face conditions. Different colors indicate positive or negative values; times indicate peak latencies or difference wave averaging intervals. Head maps are individually scaled to clearly show individual between-condition differences. (c) Representative non-latency-jitter-corrected grand-average event-related potentials (ERPs) in response to high-attractive (solid black lines), averaged (dashed black lines), and low-attractive (red lines) human faces, and chimpanzee faces (dashed red lines). The

representative ERP waveforms are collapsed across posterior left lateral (left column) and right lateral (right column) extended 10–20 scalp EEG recording locations, indicated by black dots on the peak and difference wave scalp maps (see the EEG acquisition, preprocessing, and data reduction section; recording sites shown outside the radius of the head represent locations that are below the equatorial FPZ–T7–T8–OZ plane of the assumed spherical head model). Negative polarity is oriented upward

For each individual participant and condition, we first identified the latency of the positive-polarity P1 and P2 ERP components that, respectively, preceded and followed the N170 in their averaged left- and right-hemisphere regional ERP waveforms (see the [ERP Quantification](#) section above); the P1 and P2 latencies were defined as the time points of the most positive values of the noncorrected waveform in the 75- to 150-ms and 175- to 275-ms time windows, respectively. We next searched for the largest negative peak between the P1 and P2 latencies for each trial at these sites, with the latency of this negative peak being taken as the N170 latency for a given trial. On average, the LJC procedure produced mean latency shifts of 21 ± 17 ms per condition per electrode, with the latency shifts ranging from 0 to ± 76 ms. To determine that this peak-picking technique was detecting true single-trial N170 responses rather than noise, we

applied a chi-square test for goodness of fit, testing the observed single-trial latency distribution against the null hypothesis of a uniform distribution (Spencer et al. 2000). If the peak-picking technique detected random noise, then we would expect that the distribution of single-trial negative peak latencies would be uniformly distributed across the P1-to-P2-component interval. However, if the peak-picking technique reliably detected true single-trial N170 responses, then the latencies of these responses should be clustered around a mean latency and their distribution should deviate from uniformity. The chi-square tests were significant for each participant, condition, and electrode of interest (all $ps < .05$, after correction for multiple comparisons across electrodes via the Holm–Bonferroni procedure). Hence, we concluded that the peak-picking technique was successfully detecting single-trial N170 responses and not noise.

After identifying the single-trial N170 latencies, we computed latency-jitter-corrected (LJC) waveforms at electrode sites of interest by shifting the N170 peak on each trial in time so that epochs were aligned to the mean single-trial N170 latency. We then quantified the mean amplitudes of the LJC N170 responses in the same manner as the amplitudes of the non-LJC ERPs. In addition, we used the standard deviation of the single-trial N170 latencies as a measure of latency jitter for each participant, condition, and electrode of interest. We analyzed the mean LJC N170 amplitudes, single-trial latencies, and latency standard deviations via univariate ANOVAs, with all tests corrected for nonsphericity and for multiple comparisons.

RT/N170/face rating regression analysis

We examined the general within-condition predictive relationship between processing fluency and the N170 response by regressing RTs against N170 amplitude across participants after collapsing across human face conditions; a separate analysis of the chimpanzee responses was also performed. We also examined the across-condition predictive relationship between the N170s and RTs by first arranging the data according to the theoretically predicted monotonic N170 amplitude increase and RT decrease over the averaged, to high-attractive, to low-attractive face conditions, and then computing the slopes of the across-condition changes for each variable separately for each participant. We used linear regression to predict the slopes of the across-condition changes for the N170 from the corresponding slopes for the RTs. Finally, we examined the relative within- and across-condition contributions of processing fluency, attractiveness, and distinctiveness to the N170 response by multiply regressing N170 amplitudes against RTs and the attractiveness/distinctiveness ratings of the human faces viewed during the categorization task.

These analyses were carried out via reweighted least-trimmed-squares linear regression (Rousseeuw and Leroy 2003), a robust statistical procedure that is not unduly affected by outliers and violations of model assumptions. Least-trimmed-squares regression finds the subset k of n data points with the minimal sum of squared residuals. Here, k was determined a priori according to a 95 %-confidence-interval global exclusion threshold for outliers. On average, $6.6 \% \pm 0.4 \%$ of the data points were excluded from the final across- and within-participant tests relating RT to N170 amplitude, whereas $7.3 \% \pm 0.4 \%$ of the data points were excluded from the tests relating N170 amplitude to the attractiveness/distinctiveness ratings and RTs. We implemented the regressions via the Forward Search Data Analysis toolbox for MATLAB (Riani, Perrotta, and Torti 2012) and applied them to the mean RTs and mean latency-jitter-corrected and -noncorrected N170

amplitudes computed from the correct trials entering into the ERP averages after artifact rejection. The left- and right-hemisphere N170 responses were analyzed separately, with the p values of all regressions being corrected for multiple comparisons. We report standardized regression coefficients (β) to facilitate comparisons among the predictive conditions.

Results

Behavior

Accuracy A two-way ANOVA of the human-versus-chimpanzee categorization hit rates yielded a main effect of block group, $F(3, 141) = 8.32, p < .001, \epsilon = .77, \eta^2_p = .15$, a main effect of face type, $F(3, 141) = 58.96, p < .001, \epsilon = .35, \eta^2_p = .56$, and a Block Group \times Face Type interaction, $F(9, 423) = 6.96, p < .001, \epsilon = .40, \eta^2_p = .13$. Overall, participants classified the human faces more accurately than the chimpanzee faces (high-attractive, $99.2 \% \pm 0.2 \%$; low-attractive, $99.3 \% \pm 0.1 \%$; averaged, $99.2 \% \pm 0.2 \%$; chimpanzee, $88.9 \% \pm 1.4 \%$); no accuracy difference emerged among the human faces, $p_s > .984$. However, accuracy in classifying the chimpanzee faces decreased across block groups [Block Group 1, $91.4 \% \pm 1.3 \%$; Block Group 4, $86.3 \% \pm 1.9 \%$; $F(3, 141) = 8.95, p < .001, \epsilon = .78, \eta^2_p = .16$, post-hoc $p < .006$], while remaining constant for the human faces, $p_s > .296$.

Reaction time A two-way ANOVA of the human-versus-chimpanzee categorization RTs yielded a main effect of block group, $F(3, 141) = 15.52, p < .001, \epsilon = .72, \eta^2_p = .25$, a main effect of face type, $F(3, 141) = 213.18, p < .001, \epsilon = .41, \eta^2_p = .82$, and a Block Group \times Face Type interaction, $F(9, 423) = 3.71, p < .001, \epsilon = .86, \eta^2_p = .07$. Overall, participants were faster in classifying human than chimpanzee faces (high-attractive, 383 ± 9 ms; low-attractive, 388 ± 9 ms; averaged, 384 ± 9 ms; chimpanzee, 450 ± 9 ms). Importantly, participants categorized the high-attractive and averaged faces more rapidly than the low-attractive faces, post-hoc $p_s < .009$. Furthermore, the RTs to categorize human faces significantly decreased over block groups (Block Group 1, 401 ± 10 ms; Block Group 4, 370 ± 9 ms; RT values collapsed across human face categories), $F(3, 141) = 19.45, p < .001, \epsilon = .71, \eta^2_p = .29$, post-hoc $p < .006$; a decrease in RTs to classify the chimpanzee faces only approached significance, $F(3, 141) = 2.83, p < .053, \epsilon = .81, \eta^2_p = .06$.

Attractiveness and distinctiveness ratings

Attractiveness A one-way ANOVA of the human face attractiveness ratings gathered after the species categorization task revealed a significant main effect of face type, $F(2, 94) =$

336.42, $p < .001$, $\varepsilon = .98$, $\eta^2_p = .88$. Participants rated the high-attractive faces (5.17 ± 0.12) and averaged faces (5.70 ± 0.15) as being more attractive than the low-attractive faces (2.02 ± 0.09 , post-hoc $ps < .003$). In addition, the averaged faces were rated as being more attractive than the high-attractive faces (post-hoc $p < .003$). These findings replicate previous observations that averaged faces are judged to be highly attractive (Bronstad et al. 2008; Langlois and Roggman 1990).

Distinctiveness A one-way ANOVA of the distinctiveness ratings revealed a significant main effect of distinctiveness, $F(2, 94) = 16.27$, $p < .001$, $\varepsilon = .95$, $\eta^2_p = .26$. The averaged faces (3.52 ± 0.16) were rated as being less distinctive than either the high-attractive (4.27 ± 0.17) or the low-attractive (4.84 ± 0.18) faces (post-hoc $ps < .009$). The high-attractive faces were rated as being less distinctive than the low-attractive faces, post-hoc $p < .025$.

Predictive relationship between attractiveness and distinctiveness ratings The generalized estimating equation-based regression showed that the attractiveness ratings were significantly associated with the distinctiveness ratings, $B = -0.23$, Wald $\chi^2(1, N = 48) = 16.01$, $p < .001$: In general, the more attractive the faces were rated, the less distinctive they were rated.

Event-related potentials

Figure 2 shows the grand-average ERP scalp topographies (panel A), difference topographies (panel B), and representative waveforms for each condition over the left and right hemispheres (panel C). The N170 showed the typical lateral occipitotemporal distribution of this component, as well as a positive counterfield over the scalp vertex called the *vertex positive potential* (Jeffreys 1989, 1996), which reflects volume conduction of electric signals emanating from the same neural generators as the N170 (Joyce and Rossion 2005). The N170 differences were mostly restricted to the lateral scalp areas, except for the contrasts between the chimpanzee and human faces, which also showed vertex positive potential differences over the medial central posterior scalp (these differences were not assessed).

Non-LJC analysis ANOVAs of the noncorrected N170 amplitudes yielded a significant main effect of face type, $F(3, 141) = 37.88$, $p < .001$, $\varepsilon = .74$, $\eta^2_p = .45$. The main effect of hemisphere was marginally significant, $F(1, 47) = 3.96$, $p < .052$, $\varepsilon = 1.00$, $\eta^2_p = .08$, but the Face Type \times Hemisphere interaction was not significant, $F(3, 141) = 1.33$, $p < .269$, $\varepsilon = .70$, $\eta^2_p = .03$. Follow-up analyses of the face type main effect showed that the N170 was significantly larger for the low-attractive human faces ($-3.29 \pm 0.25 \mu\text{V}$) than for the high-attractive ($-3.07 \pm 0.25 \mu\text{V}$) and averaged ($-3.08 \pm 0.24 \mu\text{V}$) human faces, post-hoc $ps < .048$, but the latter two face categories did not differ from one another, post-

hoc $ps < .909$. In addition, the N170 response to chimpanzee faces ($-4.18 \pm 0.28 \mu\text{V}$) was significantly larger than those to all the human faces, post-hoc $ps < .006$.

Analysis of the N170 latencies yielded a significant main effect of face type, $F(3, 141) = 44.75$, $p < .001$, $\varepsilon = .80$, $\eta^2_p = .49$, which indicated that N170 responses were delayed for chimpanzee faces ($179 \pm 1 \text{ ms}$) relative to the human faces (high-attractive, $170 \pm 1 \text{ ms}$; averaged, $171 \pm 2 \text{ ms}$; low-attractive, $171 \pm 2 \text{ ms}$; post-hoc $ps < .006$). We found no N170 latency differences among the human faces, post-hoc $ps > .406$. In addition, a main effect of hemisphere, $F(1, 47) = 4.91$, $p < .032$, $\varepsilon = 1.00$, $\eta^2_p = .10$, indicated that the overall N170 responses were delayed over the left ($174 \pm 2 \text{ ms}$) versus the right ($171 \pm 1 \text{ ms}$) hemisphere.

LJC analysis The analysis of the N170 amplitudes after latency jitter correction yielded effects similar to those in the noncorrected analysis. The main effect of face type was significant, $F(3, 141) = 49.24$, $p < .001$, $\varepsilon = .55$, $\eta^2_p = .51$, but the main effect of hemisphere, $F(1, 47) = 2.23$, $p < .142$, $\varepsilon = 1.00$, $\eta^2_p = .05$, and the Face Type \times Electrode Location interaction, $F(3, 141) = .24$, $p < .802$, $\varepsilon = .71$, $\eta^2_p = .01$, were not. Follow-up analyses of the face type main effect showed that the LJC N170 was significantly larger for the low-attractive human faces ($-8.05 \pm 0.29 \mu\text{V}$) than for the high-attractive ($-7.70 \pm 0.27 \mu\text{V}$) and averaged ($-7.84 \pm 0.27 \mu\text{V}$) human faces, post-hoc $ps < .015$, but the latter two face categories did not differ from one another, post-hoc $ps < .114$. In addition, the LJC N170 response to chimpanzee faces ($-9.18 \pm 0.34 \mu\text{V}$) was significantly larger than those to all of the human faces, post-hoc $ps < .006$.

The ANOVA of single-trial latency jitter yielded a main effect of face type, $F(3, 141) = 6.39$, $p < .001$, $\varepsilon = .85$, $\eta^2_p = .12$, which indicated that LJC N170 latencies were more variable for chimpanzee ($28 \pm 1 \text{ ms}$) than for the human faces (high-attractive, $24 \pm 1 \text{ ms}$; averaged, $25 \pm 1 \text{ ms}$; low-attractive, $25 \pm 1 \text{ ms}$; post-hoc $ps < .048$). We found no latency jitter differences among the human faces, post-hoc $ps > .802$.

RT/N170 regression analysis In general, larger N170 responses were associated with slower RTs within conditions. This was the case for the noncorrected N170 over the left scalp hemisphere within the collapsed human face condition [$\beta_{\text{left}} = -.34 \pm .14$, $t(45) = -2.39$, $p < .042$] and over both hemispheres within the chimpanzee face condition [$\beta_{\text{left}} = -.20 \pm .07$, $t(42) = -2.83$, $p < .014$; $\beta_{\text{right}} = -.14 \pm .06 \text{ ms}/\mu\text{V}$, $t(44) = -2.11$, $p < .041$]. After N170 latency jitter correction, these within-condition N170/RT relationships were only significant at left hemisphere sites within both conditions [human faces, $\beta_{\text{left}} = -.40 \pm .14$, $t(45) = -2.86$, $p < .013$; chimpanzee faces, $\beta_{\text{left}} = -.33 \pm .14$, $t(45) = -2.36$, $p < .046$]. Furthermore, increases in N170 amplitudes (negative slopes) across the human face conditions (from averaged, to high-attractive, to

low-attractive) were significantly associated with RT increases (positive slopes) across these conditions. This was the case for both the noncorrected and corrected N170s at both left- and right-hemisphere scalp locations [noncorrected: $\beta_{\text{left}} = -.49 \pm .13$, $t(44) = -3.73$, $p < .001$; $\beta_{\text{right}} = -.33 \pm .15$, $t(44) = -2.24$, $p < .03$; corrected: $\beta_{\text{left}} = -.44 \pm .14$, $t(44) = -3.17$, $p < .006$; $\beta_{\text{right}} = -.34 \pm .15$, $t(43) = -2.29$, $p < .027$].

N170/RT/human face ratings multiple regression analysis Across experimental conditions, RTs were significantly associated with left-hemisphere N170 amplitudes within the context of the additional predictors of the attractiveness and distinctiveness ratings of the human faces; this was the case both before and after latency jitter correction of the N170s [noncorrected, $\beta_{\text{left}} = -.52 \pm .13$, $t(42) = -3.97$, $p < .002$; corrected, $\beta_{\text{left}} = -.42 \pm .12$, $t(43) = -3.39$, $p < .01$]. Furthermore, decreases in ratings of attractiveness across the human face conditions (negative slopes) were significantly associated with left-hemisphere N170 amplitude increases (negative slopes), but only after N170 latency jitter correction [$\beta_{\text{left}} = .45 \pm .13$, $t(43) = 3.54$, $p < .006$]. In contrast, the distinctiveness ratings were not significantly associated with N170 amplitudes across conditions, either before or after latency jitter correction ($p = 1$). We ascertained whether the attractiveness ratings and/or the RTs were associated with the N170 responses across conditions beyond the contribution of the other variable (plus distinctiveness held constant). We achieved this via use of the standard R^2 change statistic, applied to the regression outcome for the left-hemisphere LJC N170 amplitudes. We found that attractiveness ratings and RTs were both associated with N170 amplitude, beyond the contribution of the other variable, to roughly equal degrees [attractiveness, $R^2_{\text{change}} = .25$, $F(1, 41) = 16.98$, $p < .001$; RT, $R^2_{\text{change}} = .20$, $F(1, 41) = 13.52$, $p < .001$]. Finally, none of the three variables were significantly associated with N170 amplitudes within conditions ($p < .32$).

Discussion

The present study has provided a test of the averageness theory of facial attractiveness by comparing behavioral and ERP responses to high-attractive, averaged, and low-attractive faces presented in the context of a simple species face categorization task. To our knowledge, this is the first study to directly compare neurobehavioral responses among facial prototypes and individual faces varying in attractiveness.

Attractive and averaged faces are more prototypical than unattractive faces

Our first finding demonstrated an inverse relationship between the attractiveness and distinctiveness of our face stimuli. Averageness theory predicts that attractive faces similar to the

prototype should be perceived as being more typical and less distinctive than unattractive faces that differ from the prototype. Consistent with this prediction, participants rated high-attractive and averaged faces as being more attractive and less distinctive than the low-attractive faces. In addition, they rated the averaged faces as being more attractive and less distinctive than the high-attractive faces, consistent with the expectation that averaged faces should be even more prototypical than high-attractive faces. Furthermore, regression analysis showed that attractiveness was negatively associated with distinctiveness—the more attractive a face was, the less distinctive was its rating. These findings replicate previous observations of an inverse relationship between facial attractiveness and facial distinctiveness (Light et al. 1981; Peskin and Newell 2004; Rhodes and Tremewan 1996; Vokey and Read 1992) and support the claim that attractive and averaged faces are more prototypical than unattractive faces.

In closing this section, we should note that faces with atypical features are sometimes rated as being highly attractive, or even more attractive than averaged faces (Alley and Cunningham 1991; Cunningham, Barbee, and Pike 1990; Mende-Siedlecki, Said, and Todorov 2013; Perrett, May, and Yoshikawa 1994; Wickham and Morris 2003). However, comparison of these findings with the predictions of averageness theory (or any other theory, for that matter) is impeded by the fact that measurement of the relationship between attractiveness and distance from the prototype varies across studies. For example, some studies measure facial typicality (the degree of deviation from an averaged face), whereas other studies measure facial distinctiveness (the degree to which a face stands out in a crowd). Wickham and Morris directly compared these two measures and found a strong linear inverse relationship between attractiveness and typicality, but a curvilinear relationship between attractiveness and distinctiveness. They explained their findings by suggesting that some high-attractive faces might be similar to an averaged face overall, yet also distinctive due to certain unique factors, such as the presence of an unusual facial feature or the similarity of a face to famous people or one's friends and relatives.⁴ Thus, the findings of Wickham and Morris support the basic claim of averageness theory by suggesting that proximity to the facial prototype is necessary for a face to be considered attractive, although such attractiveness may be modulated by other known factors, such as individual

⁴ We also suggest that attractiveness and distinctiveness ratings could be distorted by using stimulus sets that are unbalanced in terms of the numbers of attractive and unattractive faces, as was done in some of these studies (Mende-Siedlecki et al. 2013; Perrett et al. 1994). For example, a small number of attractive faces embedded in a larger set of less attractive faces could appear to be highly distinctive due to their low frequency, whereas the attractiveness ratings of a set of extremely attractive faces could be driven by distinctiveness factors, since participants would search for any criteria to distinguish among the faces. In the present study, our stimulus set was balanced among attractive and unattractive faces, so we measured the distinctiveness of the faces due to the greater generality of this property.

rater preferences, facial sexual dimorphism, symmetry, familiarity, and hormone-dependent facial features (Fink, Neave, Manning, and Grammer 2006; Hönekopp 2006; Hoss et al. 2005; Penton-Voak et al. 2001; Perrett et al. 1998; Peskin and Newell 2004; Rennels et al. 2008; Rhodes et al. 1998; Rhodes and Tremewan 1996; Rubenstein et al. 2002; Thornhill and Gangestad 1999).

Attractive and averaged faces engage fluent neurocognitive processing

Our second finding showed that participants categorized high-attractive and averaged faces as “human” faster than low-attractive faces. These RT differences, though small (~4–5 ms), were highly reliable and are consistent with a previous observation that attractiveness facilitates the speed and accuracy of gender-based face classification (Hoss et al. 2005). Importantly, the present RT findings demonstrate for the first time that, like high-attractive faces, prototypical (averaged) faces are processed more fluently than unattractive faces. Given that our participants also rated the averaged faces as being highly attractive, these findings support the premise of averageness theory—that high-attractive faces are attractive and unattractive faces are unattractive by virtue of their respective similarity and dissimilarity to the averaged face, and by virtue of the effects that this has on the ease of face processing.

Our third finding demonstrated that the more-fluent processing of averaged and high-attractive faces was accompanied by a reduced N170 ERP component relative to the N170 elicited by the less fluently processed low-attractive faces. We observed these findings both before and after latency jitter correction of the N170 (see the next section). These findings replicate a previously observed N170 reduction in response to attractive versus unattractive faces (Halit et al. 2000), while extending this observation to the case of averaged faces. Our study also extends Halit et al.’s by examining these attractiveness-related N170 changes in the context of similar changes in RTs (the earlier study used a simple passive-viewing paradigm). This allowed us to draw conclusions about the fluency of the neurocognitive processes indexed by the N170. Regression analyses showed that larger N170 amplitudes were associated with longer RTs both within and across the human face conditions, both before and after N170 latency jitter correction. In addition, RTs were associated with the noncorrected and corrected N170 amplitudes across conditions in the context of additional predictors of attractiveness and distinctiveness ratings. Our findings are also consistent with the well-known N170 face inversion effect (i.e., an enhanced N170 for inverted vs. upright photorealistic faces; Itier and Taylor 2002, 2004a, 2004b; Rossion et al. 1999; Rossion et al. 2003; Vizioli et al. 2010) that is thought to arise from the disruption of configural processing of inverted faces, which is preserved for upright faces (Rossion and Gauthier

2002). Here, we conjecture that the different levels of fluent processing do not arise from a disruption of configural processing per se, but instead are due to the configural similarity of the high-attractive and averaged faces to facial prototypes developed through life-long experience with faces.

A further observation of interest is that the regression analyses indicated that decreases in attractiveness ratings across the averaged to high- to low-attractive conditions predicted corresponding increases in LJC N170 amplitudes. However, analysis of the R^2 change statistic showed that both attractiveness and RT were associated with the LJC N170 amplitudes in roughly equal manners. That is, both behavioral indices were each uniquely related to the N170 response, and thus do not reflect exactly the same cognitive processes. We speculate that the attractiveness ratings index the effect of distance from the prototype on processing fluency during the early perceptual stage indexed by the N170, whereas RTs may reflect the effects that such fluent processing has on the later cognitive processes directly responsible for the production of an overt motor response.

We also observed poorer categorization accuracy, slower RTs, and a larger N170 for chimpanzee than for human faces, a finding consistent with the interpretation that the former were processed less fluently than the latter. However, this interpretation is confounded by the fact that the chimpanzee stimuli were not matched with the human faces in terms of physical stimulus properties or stimulus probability; infrequent stimuli, such as the chimpanzee faces, are especially well-known to produce increased RTs and ERP amplitudes, all other factors being equal (Blackman 1972; Squires, Donchin, Herning, and McCarthy 1977). Nevertheless, interpretation of the regression analysis relating the N170 to RTs within the chimpanzee face condition is not subject to these confounds, because between-condition differences in stimulus properties and frequencies were not relevant. Here, the regression analysis showed that larger latency-jitter-corrected and -noncorrected N170 amplitudes were associated with longer RTs within the chimpanzee face condition. This finding provides further evidence that the N170 reflected the processing fluency of the faces.

We note that averageness theory also predicts that, to the degree that the averaged faces were closer to the prototype than were the high-attractive individual faces, the former should be more fluently processed than the latter. This fluent processing should manifest as differences in RTs and N170 responses. However, we found no mean RT/N170 differences between the high-attractive individual and averaged faces, suggesting that both kinds of faces were processed with roughly equal fluency. We speculate that these null findings between the high-attractive and averaged faces indicate that the early stages of the visual system treat different faces as being highly similar when they are within a certain degree of proximity to the prototype. The averaged faces that we utilized were arguably closer to the prototype than the high-attractive

faces because the former were composites of 32 individual exemplars; this conclusion is also supported by the lower distinctiveness ratings for the averaged than for the high-attractive faces. However, we did not explicitly quantify how much closer the averaged faces were to the prototype than were the high-attractive faces; it may be that the distance between the two face types was small enough that only perceptual differences (i.e., attractiveness or distinctiveness), and not fluency differences, emerged between the two face categories. Nevertheless, it is likely that if we were to use slightly less attractive faces for the high-attractive face condition and/or to use a better estimate of the facial prototype (e.g., by using air brushing or mixed-race faces during the construction of the average faces), fluency differences between the averaged and individual high-attractive faces would emerge. This raises an interesting question for future research: Namely, what is the criterion distance at which different faces will be treated as being similar to one another by the brain? This question could be answered by a future study employing a finer parametric variation of attractiveness and distance from the prototype among individual and averaged faces.

Could it be the case that, rather than fluency, our N170 effects merely reflected increased salience of the high-attractive and averaged faces relative to the low-attractive faces? This is unlikely, because previous studies linking stimulus salience to the N170 have observed a larger response for this component for salient than for nonsalient stimuli (Batty and Taylor 2003; Blau, Maurer, Tottenham, and McCandliss 2007), a pattern opposite to the one observed here. It is also unlikely that our N170 effects were due to the low-attractive faces being more salient because of their high distinctiveness; salient stimuli should be more easily detectable and/or discriminable, and thus should have led to faster categorization RTs for this condition, which is again a pattern opposite to the one observed here. Furthermore, regression analysis showed that ratings of facial distinctiveness were not associated with N170 amplitudes in the present data set. Finally, these N170 effects cannot be due to the low-attractive faces possessing salient low-level stimulus properties, since these were matched across face categories (see the [supplementary materials](#)).

We close this section by noting that these N170 fluency effects are most likely not specific to faces, because this ERP component is also larger for subordinate- versus basic-level nonface object categorization (Tanaka, Luu, Weisbrod, and Kiefer 1999), as well as for nonface objects of expertise versus objects of nonexpertise (Busey and Vanderkolk 2005; Rossion, Collins, Goffaux, and Curran 2007; Tanaka and Curran 2001). This suggests a more general N170 sensitivity to entry-level object categorization (Rossion et al. 2002). Indeed, the N170 evidence implies that expert object categorization performance recruits experience-dependent early-stage neural resources that are similar to those used for faces (Busey and Vanderkolk 2005; Rossion et al. 2007; Tanaka and Curran 2001), although the

exact mechanism of how this is accomplished is still unknown (Rossion and Jacques 2011). Nonetheless, it is important to distinguish these general N170 expertise effects from the N170 fluency effects reported here. General expertise effects are defined via across-category comparisons (e.g., cars vs. birds, faces vs. nonfaces), and thus reflect a preference for objects of expertise over other objects. In contrast, the present N170 effects are defined via a comparison within a single object category (e.g., human faces), and thus reflect a preference for a particular configuration of the same object. Hence, like the N170 face inversion effect, our findings suggest that when category structure is held constant, a prototypical object will elicit a smaller N170 response than will a nonprototypical object. However, if both stimuli are also objects of expertise, they will produce a larger N170 response than do objects of nonexpertise (absent the presence of other confounding factors, such as the low stimulus frequency of the chimpanzee faces relative to the human faces in the present study). This suggestion is supported by the N170 face inversion effect, as well as by an N170 inversion effect for nonface objects of expertise (Rossion et al. 2002) and the impairment of expert nonface object categorization performance with inversion (Diamond and Carey 1986; Rossion and Curran 2010).

N170 fluency effects reflect the engagement of fewer neural resources

Our fourth finding suggests that the overall reduction in N170 differences for high-attractive and averaged faces versus low-attractive faces resulted from differences in the degrees to which these stimuli engaged neural resources, and not from a difference in the temporal consistency of that resource engagement. Our single-trial EEG latency jitter correction technique produced jitter-free measures of ERP amplitudes to index the relative contributions of amplitude and timing changes to the between-category differences that were evident in the noncorrected ERPs. Corrected N170 amplitudes were smaller for high-attractive and averaged faces than for low-attractive faces, and were significantly associated with longer RTs within and across face conditions. Nevertheless, no single-trial EEG latency jitter difference accompanied these N170 amplitude differences. This indicates that the N170 differences were due to differences in the amplitudes of the neural activity underlying the N170. Additional support for this conclusion is that we did not observe any N170 latency delays for one condition relative to another, in either the conventional ERP analysis or the LJC analysis. Such delays (~10–20 ms) are often observed for inverted versus upright faces (Bentin et al. 1996; Rossion et al. 1999; Rossion et al. 2000) or when facial configuration information is removed or degraded (as happens with inversion). These latency delays arise from either a lag in the onset of face processing or a slower rate at which face information is accumulated by the neurons representing the faces (Rossion and

Jacques 2011). Such latency delays should not be expected in the present data, since all of our stimuli portrayed complete facial configurations in upright orientations. The fact that we observed no such N170 delays provides additional evidence that timing differences were not a factor in the present data.

Our finding of an overall reduction in latency jitter corrected N170 amplitude to high-attractive and averaged faces versus low-attractive faces is consistent with previous observations of decreased hemodynamic activity in human visual cortex for prototypical versus nonprototypical faces and visual dot patterns (Loffler et al. 2005; Reber et al. 1998) and with larger firing rates of face-selective neurons in the inferotemporal cortex of nonhuman primates for nonprototypical individual faces versus the average of the faces (Leopold et al. 2006). Our findings are also consistent with the conclusion of Rousselet, Husk, Bennett, and Sekuler (2007) that differences in the timing of single-trial EEG responses do not contribute to the large N170 face effect (i.e., larger N170s for faces than for objects). Rousselet et al. showed that the N170 face effect arises from between-condition EEG amplitude differences in the 5- to 15-Hz range, via wavelet-based time-frequency analysis, which can decompose an EEG signal into the component signals of individual frequencies in order to separately quantify their amplitudes and phases. However, this prior study was unable to reach a definite conclusion regarding the contribution of single-trial EEG timing changes to the much smaller N170 face inversion effect, because of the poor temporal resolution of wavelet analysis at low frequencies. In contrast, the single-trial analysis used here does not suffer from this problem, because it directly quantified single-trial peak latencies on each trial with millisecond resolution. Thus, we were able to reach a decisive conclusion about the contributions of amplitude and timing changes to the even smaller attractiveness-related N170 effects observed here. One suggestion for future research is to apply our single-trial analysis method to studies of the face inversion effect to confirm the likelihood that this effect also depends purely on between-condition changes in EEG amplitudes.

Conclusion

We have provided novel evidence that high-attractive and mathematically averaged faces both engage fluent facial processing at early stages of visual perception. The present findings, together with previous observations (Bronstad et al. 2008; Langlois and Roggman 1990; Langlois et al. 1994; Rhodes and Tremewan 1996; Rubenstein et al. 1999), support the theory that faces are perceived as being attractive when they approximate an average facial configuration.

Author note We thank Arian Mobasser, Stevie Schein, and Erin Anderson for assistance with the data collection. We thank Connor Principe and three anonymous reviewers for comments on earlier drafts of the manuscript. This research was supported by the National Institutes

of Health [Grant No. HD021332 to J.H.L.; Grant No. 5T32MH065728-08 through the Texas Consortium in Behavioral Neuroscience Diversity Training Program to L.T.T.].

References

- Alley, T. R., & Cunningham, M. R. (1991). Averaged faces are attractive, but very attractive faces are not average. *Psychological Science*, *2*, 123–125.
- Apicella, C. L., Little, A. C., & Marlowe, F. W. (2007). Facial averageness and attractiveness in an isolated population of hunter-gatherers. *Perception*, *36*, 1813–1820.
- Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, *17*, 613–620.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, C. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.
- Blackman, A. R. (1972). Stimulus probability and choice reaction time. *Perception & Psychophysics*, *12*, 146–150.
- Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions*, *3*, 7. doi:10.1186/1744-9081-3-7
- Bronstad, P. M., Langlois, J. H., & Russell, R. (2008). Computational models of facial attractiveness judgments. *Perception*, *37*, 126–142.
- Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research*, *45*, 431–448.
- Cunningham, M. R., Barbee, A. P., & Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness. *Journal of Personality and Social Psychology*, *59*, 61–72.
- Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., & Lueschow, A. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *NeuroImage*, *35*, 1495–1501. doi:10.1016/j.neuroimage.2007.01.034
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. doi:10.1016/j.jneumeth.2003.10.009
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117. doi:10.1037/0096-3445.115.2.107
- Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel scenes. *Journal of Cognitive Neuroscience*, *13*, 171–180.
- Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, *11*, 154–158. doi:10.1111/1467-8721.00190
- Fink, B., Neave, N., Manning, J. T., & Grammer, K. (2006). Facial symmetry and judgments of attractiveness, health and personality. *Personality and Individual Differences*, *41*, 491–499.
- Gardiner, J. C., Luo, Z., & Roman, L. A. (2009). Fixed effects, random effects and GEE: What are the differences? *Statistics in Medicine*, *28*, 221–239. doi:10.1002/sim.3478
- Ghisletta, P., & Spini, D. (2004). An introduction to generalized estimating equations and an application to assess selectivity effects in a longitudinal study on very old individuals. *Journal of Educational and Behavioral Statistics*, *29*, 421–437.
- Gratton, G., Kramer, A. F., Coles, M. G. H., & Donchin, E. (1989). Simulation studies of latency measures of components of the event-related brain potential. *Psychophysiology*, *26*, 233–248.
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112. doi:10.1007/BF02289823

- Griffin, A. M., & Langlois, J. H. (2006). Stereotype directionality and attractiveness stereotyping: Is beauty good or is ugly bad? *Social Cognition*, *24*, 187–206.
- Gruber, W. R., Klimesch, W., Sauseng, P., & Doppelmayr, M. (2005). Alpha phase synchronization predicts P1 and N1 latency and amplitude size. *Cerebral Cortex*, *15*, 371–377.
- Halgren, E., Raji, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex*, *10*, 69–81.
- Halit, H., de Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *NeuroReport*, *11*, 1871–1875.
- Hamermesh, D. S. (2011). *Beauty pays: Why attractive people are more successful*. Princeton: Princeton University Press.
- Harmon-Jones, E., & Allen, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual difference approaches. *Personality and Social Psychology Bulletin*, *27*, 889–898.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233. doi:10.1016/S1364-6613(00)01482-0
- He, P., Wilson, G., & Russell, C. (2004). Removal of ocular artifacts from electro-encephalogram by adaptive filtering. *Medical and Biological Engineering and Computing*, *42*, 407–412.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.
- Hönekopp, J. (2006). Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 199–209. doi:10.1037/0096-1523.32.2.199
- Hoss, R. A., Ramsey, J. L., Griffin, A. M., & Langlois, J. H. (2005). The role of facial attractiveness and facial masculinity/femininity in sex classification of faces. *Perception*, *34*, 1459–1474.
- Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *NeuroImage*, *15*, 353–373.
- Itier, R. J., & Taylor, M. J. (2004a). Effects of repetition learning on upright, inverted, and contrast-reversed face processing using ERPs. *NeuroImage*, *21*, 1518–1532.
- Itier, R. J., & Taylor, M. J. (2004b). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex*, *14*, 132–142.
- Jacques, C., & Rossion, B. (2007). Early electrophysiological responses to multiple face orientations correlate with individual discrimination performance in humans. *NeuroImage*, *36*, 863–876.
- Jáskowski, P., & Verleger, R. (2000). An evaluation of methods for single-trial estimation of P3 latency. *Psychophysiology*, *37*, 153–162. doi:10.1111/1469-8986.3720153
- Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research*, *78*, 193–202.
- Jeffreys, D. A. (1996). Evoked studies of face and object processing. *Visual Cognition*, *6*, 1–38.
- Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology*, *116*, 2613–2631.
- Judge, T. A., Hurst, C., & Simon, L. S. (2009). Does it pay to be smart, attractive, or confident (or all three)? Relationships among general mental ability, physical attractiveness, core self-evaluations, and income. *Journal of Applied Psychology*, *94*, 742–755. doi:10.1037/a0015497
- Junghöfer, M., Elbert, T., Tucker, D. M., & Braun, C. (1999). The polar average reference effect: A bias in estimating the head surface integral in EEG recording. *Clinical Neurophysiology*, *110*, 1149–1155.
- Kalick, S. M., Zebrowitz, L. A., Langlois, J. H., & Johnson, R. M. (1998). Does human facial attractiveness honestly advertise health? *Psychological Science*, *9*, 8–13.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.
- Klimesch, W., Schack, B., Schabus, M., Doppelmayr, M., Gruber, W., & Sauseng, P. (2004). Phase-locked alpha and theta oscillations generate the P1–N1 complex and are related to memory performance. *Cognitive Brain Research*, *19*, 302–316.
- Komatsu, L. K. (1992). Recent views of conceptual structure. *Psychological Bulletin*, *112*, 500–526. doi:10.1037/0033-2909.112.3.500
- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, *1*, 115–121.
- Langlois, J. H., Roggman, L. A., Casey, R. J., & Ritter, J. M. (1987). Infant preferences for attractive face: Rudiments of a stereotype? *Developmental Psychology*, *23*, 363–369.
- Langlois, J. H., Roggman, L. A., & Rieser-Danner, L. A. (1990). Infants' differential social responses to attractive and unattractive faces. *Developmental Psychology*, *26*, 153–159.
- Langlois, J. H., Roggman, L. A., & Musselman, L. (1994). What is average and what is not average about attractive faces? *Psychological Science*, *5*, 214–220.
- Langlois, J. H., Ritter, J. M., Casey, R. J., & Sawin, D. B. (1995). Infant attractiveness predicts maternal behaviors and attitudes. *Developmental Psychology*, *31*, 464–472.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*, 390–423. doi:10.1037/0033-2909.126.3.390
- Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature*, *442*, 572–575. doi:10.1038/nature04951
- Light, L. L., Hollander, S., & Kayra-Stuart, F. (1981). Why attractive people are harder to remember. *Personality and Social Psychology Bulletin*, *7*, 269–276.
- Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., & Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neuroscience Letters*, *253*, 147–150.
- Loffler, G., Yourganov, G., Wilkinson, F., & Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience*, *8*, 1386–1390.
- Makeig, S., Westerfield, M., Jung, T. P., Enghoff, S., Townsend, J., & Courchesne, E. (2002). Dynamic brain sources of visual evoked responses. *Science*, *295*, 690–694.
- 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. doi:10.1093/scan/nsr090
- Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B., Burt, D. M., & Perrett, D. I. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proceedings of the Royal Society B*, *268*, 1617–1623.
- Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape judgments of female attractiveness. *Nature*, *368*, 239–242.
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, *394*, 884–887. doi:10.1038/29772
- Perrin, F., Perrier, J., Bertrand, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and Clinical Neurophysiology*, *66*, 75–81.

- Peskin, M., & Newell, F. N. (2004). Familiarity breeds attraction: Effects of exposure on the attractiveness of typical and distinctive faces. *Perception, 33*, 147–157.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*, 127–152. doi:10.1111/1469-8986.3720127
- Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83*, 304–308.
- Principe, C. P., & Langlois, J. H. (2011). Faces differing in attractiveness elicit corresponding affective responses. *Cognition and Emotion, 25*, 140–148. doi:10.1080/02699931003612098
- Principe, C. P., & Langlois, J. H. (2012). Shifting the prototype: Experience with faces influence affective and attractiveness preferences. *Social Cognition, 30*, 109–120. doi:10.1521/soco.2012.30.1.109
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*, 510–532. doi:10.1037/0033-2909.114.3.510
- Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998). Cortical areas supporting category learning identified using functional MRI. *Proceedings of the National Academy of Sciences, 95*, 747–750.
- Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3*, 382–407.
- Rennels, J. L., Bronstad, P. M., & Langlois, J. H. (2008). Are attractive men's faces masculine or feminine? The importance of type of facial stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 884–893. doi:10.1037/0096-1523.34.4.884
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.
- Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science, 7*, 105–110.
- Rhodes, G., Proffitt, F., Grady, J. M., & Sumich, A. (1998). Facial symmetry and the perception of beauty. *Psychonomic Bulletin & Review, 5*, 659–669. doi:10.3758/BF03208842
- Riani, M., Perrotta, D., & Torti, F. (2012). FSDA: A MATLAB toolbox for robust analysis and interactive data exploration. *Chemometrics and Intelligent Laboratory Design, 116*, 17–32.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439. doi:10.1016/0010-0285(76)90013-X
- Rossion, B., & Curran, T. (2010). Visual expertise with pictures of cars correlates with RT magnitude of the car inversion effect. *Perception, 39*, 173–183. doi:10.1068/p6270
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews, 1*, 63–75.
- Rossion, B., & Jacques, C. (2011). The N170: Understanding the time course of face perception in the human brain. In S. Luck & E. Kappenman (Eds.), *The Oxford handbook of ERP components* (pp. 115–142). New York: Oxford University Press.
- Rossion, B., Delvenne, J.-F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guérit, J.-M. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology, 50*, 173–189.
- Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not inverted objects: An electrophysiological account of face-specific processes in the human brain. *NeuroReport, 11*, 69–74.
- Rossion, B., Curran, T., & Gauthier, I. (2002). A defense of the subordinate-level expertise account for the N170 component. *Cognition, 85*, 189–196.
- Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orienting tuning for face, word, and object processing in the visual cortex. *NeuroImage, 20*, 1609–1624.
- Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization processes. *Journal of Cognitive Neuroscience, 19*, 543–555. doi:10.1162/jocn.2007.19.3.543
- Rousseeuw, P. J., & Leroy, A. M. (2003). *Robust regression and outlier detection*. Hoboken: Wiley.
- Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2007). Single-trial EEG dynamics of object and face processing. *NeuroImage, 36*, 843–862.
- Rubenstein, A. J., Kalakanis, L., & Langlois, J. H. (1999). Infant preferences for attractive faces: A cognitive explanation. *Developmental Psychology, 35*, 848–855.
- Rubenstein, A. J., Langlois, J. H., & Roggman, L. A. (2002). What makes a face attractive and why: The role of averageness in defining facial beauty. In G. Rhodes & L. A. Zebrowitz (Eds.), *Facial attractiveness: Evolutionary, cognitive, and social perspectives* (pp. 1–33). Westport: Ablex Publishing.
- Schacht, A., Werheid, K., & Sommer, W. (2008). The appraisal of facial beauty is rapid but not mandatory. *Cognitive, Affective, & Behavioral Neuroscience, 8*, 132–142. doi:10.3758/CABN.8.2.132
- Smulders, F. T. Y., Kenemans, J. L., & Kok, A. (1994). A comparison of different methods for estimating single-trial P300 latencies. *Electroencephalography and Clinical Neurophysiology, 92*, 107–114.
- Spencer, K. M. (2005). Averaging, detection, and classification of single-trial ERPs. In T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 209–228). Cambridge: MIT Press.
- Spencer, K. M., Abad, E. V., & Donchin, E. (2000). On the search for the neurophysiological manifestation of recollective experience. *Psychophysiology, 37*, 494–506.
- Squires, K. C., Donchin, E., Herning, R. I., & McCarthy, G. (1977). On the influence of task relevance and stimulus probability on event-related-potential components. *Electroencephalography and Clinical Neurophysiology, 42*, 1–14.
- Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science, 12*, 43–47.
- Tanaka, J. W., Luu, P., Weisbrod, M., & Kiefer, M. (1999). Tracking the time course of object categorization using event-related potentials. *NeuroReport, 10*, 829–835.
- Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human Nature, 4*, 237–269.
- Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences, 3*, 452–460.
- Trujillo, L. T., & Allen, J. J. B. (2007). Theta EEG dynamics of the error-related negativity. *Clinical Neurophysiology, 118*, 645–668. doi:10.1016/j.clinph.2006.11.009
- Trujillo, L. T., Kornguth, S., & Schnyer, D. M. (2009). An ERP examination of the differential effects of sleep deprivation on exogenously cued and endogenously cued attention. *Sleep, 32*, 1285–1297.
- Vizioli, L., Foreman, K., Rousselet, G. A., & Caldara, R. (2010). Inverting faces elicits sensitivity to race on the N170 component: A cross-cultural study. *Journal of Vision, 10*(1):15, 1–23. doi:10.1167/10.1.15
- Vokey, J. R., & Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition, 20*, 291–302. doi:10.3758/BF03199666

- Wickham, L. H. V., & Morris, P. E. (2003). Attractiveness, distinctiveness, and recognition of faces: Attractive faces can be typical or distinctive but are not better recognized. *American Journal of Psychology*, *116*, 455–468.
- Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000. doi:10.1037/0022-3514.81.6.989
- Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science*, *17*, 799–806. doi:10.1111/j.1467-9280.2006.01785.x
- Yao, D. (2001). A method to standardize a reference of scalp EEG recordings to a point at infinity. *Physiological Measurement*, *22*, 693–711.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.