OXFORD

The iterative nature of person construal: Evidence from event-related potentials

Hannah I. Volpert-Esmond, Edgar C. Merkle, and Bruce D. Bartholow

Department of Psychological Sciences, University of Missouri, Columbia, MO 65211, USA

Correspondence should be addressed to Hannah I. Volpert-Esmond, Department of Psychological Sciences, University of Missouri, 124 Psychology Building, Columbia, MO 65211, USA. E-mail: hannah.volpert@mail.missouri.edu.

Abstract

Recently, a dynamic-interactive model of person construal (DI model) has been proposed, whereby the social categories a person represents are determined on the basis of an iterative integration of bottom-up and top-down influences. The current study sought to test this model by leveraging the high temporal resolution of event-related brain potentials (ERPs) as 65 participants viewed male faces that varied by race (White *vs* Black), fixating either between the eyes or on the forehead. Within face presentations, the effect of fixation, meant to vary bottom-up visual input, initially was large but decreased across early latency neural responses identified by a principal components analysis (PCA). In contrast, the effect of race, reflecting a combination of top-down and bottom-up factors, initially was small but increased across early latency principal components. These patterns support the DI model prediction that bottom-up and top-down processes are iteratively integrated to arrive at a stable construal within 230 ms. Additionally, exploratory multilevel modeling of single trial ERP responses representing a component linked to outgroup categorization (the P2) suggests change in effects of the manipulations over the course of the experiment. Implications of the findings for the DI model are considered.

Key words: ERPs; multilevel modeling; race; face perception; principal components analysis

Introduction

Traditional models of person perception hold that, upon seeing a person, perceivers rely on visual information to place him or her into a relevant social category (e.g. male or female) (Fiske and Neuberg, 1990; Bodenhausen and Macrae, 1998). Activation of social categories is assumed to be automatic, supported by evidence from a variety of behavioral paradigms showing that the activation of category-related information occurs even when perceivers are under cognitive load (Macrae *et al*., 1994); when categories are irrelevant to the participant's task (Fazio *et al*., 1995); and when category-related primes are presented subliminally (Devine, 1989; Lepore and Brown, 1997). Activation of social categories subsequently impacts a number of down-stream consequences, including stereotype activation (Hehman *et al*., 2013), evaluative associations (Livingston and Brewer,

2002), non-verbal behavior (Dovidio *et al*., 1997) and trust (Stanley *et al*., 2011).

Recently, research on person construal has focused on the antecedents of social categorization rather than its consequences (Kawakami *et al*., 2017). In particular, the dynamic interactive theory of person construal (DI Model) proposes a more complex process whereby categorization decisions are not solely dictated by the visual information being perceived (Fodor, 1983), but rather reflect integration of bottom-up and top-down processes (Freeman and Ambady, 2011; Freeman *et al*., 2011). This idea incorporates knowledge about the organization of neural networks that allow for top-down inputs on primary visual cortical areas (Di Russo *et al*., 2003; Collins and Olson, 2014; Vetter and Newen, 2014; Teufel and Nanay, 2017) and the bidirectional interplay between cognition and perception (Gilbert and Li, 2013).

The DI model suggests that, when faces are the objects of perception, perceptual cues in target faces partially activate multiple competing social categories, which resolve over iterations that cycle information through higher-order and lower-order systems to arrive at a stable representation (Freeman *et al.*, 2011; Stolier and Freeman, 2016a,b). According to this model, the active representation of the face (for example, whether the person is White or Black) is initially informed largely by bottom-up processes operating on information in visual cortex and primarily reflecting objective sensory information, such as skin tone and hair texture. Subsequently, this initial, tentative representation activates higher-order neural systems that access learned information, such as stereotypes, expectations, motivations and goals of the perceiver, which then influence the active representation in a top-down manner. For example, seeing a racially ambiguous person in a business suit versus a janitor's uniform changes the likelihood that he or she will be categorized as Black or White because of learned associations between social status and race (Freeman *et al.*, 2011).

This integrative process provides a mechanism by which top-down variables can influence early social categorization processes, consistent with findings from a number of recent studies. For example, conditions of economic scarcity (Rodeheffer *et al.*, 2012; Ho *et al.*, 2013; Krosch and Amodio, 2014), political orientation (Krosch *et al.*, 2013), semantic labels (Tskhay and Rule, 2015) and motivation to be unbiased (Chen *et al.*, 2014) all have been shown to affect the categorization of racially ambiguous faces.

The current study expands on this prior work in several ways. Whereas many previous studies have used faces morphed along a racial continuum to vary the visual information they convey (Krosch and Amodio, 2014), the current study included a visual fixation manipulation to change the bottom-up influence of visual information without changing the stimuli themselves. Previous research has shown that varying fixation is effective in changing the extent to which racial category information is extracted from faces (Hills and Lewis, 2006, 2011). Here, fixation location varied between the eyes and the forehead. Fixating between the eyes is the default in spontaneous face processing (Kawakami *et al.*, 2014; Peterson and Kanwisher, 2015), and therefore is thought to convey more category-relevant information (Hills and Lewis, 2006). In contrast, the forehead is an unusual fixation location that conveys little category-relevant information. In this way, initial attention to sensory information could be manipulated without altering the faces, thereby facilitating examination of the effect of bottom-up processes.

Social category information also was manipulated by presenting faces that varied by race. Perceiving race involves both bottom-up processes, including differences in brightness and contrast related to skin tone and spatial frequencies reflecting variability in facial physiognomy (Hayward *et al.*, 2008; Zhao and Bentin, 2011), and top-down processes including accessing learned information that associates differences in facial features with distinct racial categories (Levin and Banaji, 2006). The influence of top-down processes in social categorization is analogous to the way learning and verbal labels encourage the perception of a continuous band of light frequencies as separable colors (Collins and Olson, 2014). Here, bottom-up differences were minimized by converting the images to gray scale and adjusting luminance. While not a pure distinction, incorporating experimental manipulations that differentially rely on bottom-up and top-down processes allowed us to examine the

time course of their integration. In accordance with the DI model (Freeman and Ambady, 2011), we expected the effect of fixation (mainly representing differences in bottom-up processes) to be large upon initial perception of a face but to decrease as person construal continued, whereas the effect of race (representing differences in both top-down and bottom-up processes) was expected to be small initially but to increase as processing iterations unfold.

Event-related brain potentials (ERPs) were recorded to allow observation of this theorized integration over time [see Amodio *et al.*, (2014), for background on the ERP approach]. Two methods of analyzing ERP data were used to test hypotheses derived from the DI model: (1) a traditional approach examining mean amplitude of a scalp-recorded component previously associated with social categorization (the P2 or P200; Ito and Bartholow, 2009), and (2) a principal components analysis (PCA) approach examining a sequence of underlying components contributing to early face processing.

The P2 generally peaks 150–250 ms post-stimulus along the scalp midline and has been associated with early orienting of attention to threatening or distinctive stimuli (Correll *et al.*, 2006; Kubota and Ito, 2007). Outgroup faces consistently elicit larger P2s than ingroup faces (Willadsen-Jensen and Ito, 2006, 2008; Amodio, 2010; Dickter and Kittel, 2012). This occurs regardless of task relevance (Ito and Urland, 2003, 2005; Kubota and Ito, 2007; He *et al.*, 2009) or context (Correll *et al.*, 2006; Dickter and Bartholow, 2007; Willadsen-Jensen and Ito, 2008), consistent with the notion that ingroup–outgroup distinctions occur spontaneously. Importantly, prior research indicates the P2 is sensitive to category distinctions, not simply to low-level perceptual features of faces. Specifically, Dickter and Bartholow (2007) found that while Black faces elicited larger P2 amplitude than White faces among White participants, the opposite pattern emerged among Black participants. We expected to replicate this well-established effect, such that Black faces elicit larger P2s than White faces in a predominantly White sample.

A concern with the traditional measurement of the P2 as mean amplitude within a particular time window is that it effectively removes the inherently multivariate nature of the ERP, eliminating its main advantages—its millisecond-level temporal resolution and continuous measurement over time. Therefore, we also used PCA to investigate predicted changes in the effects of our manipulations over a sequence of quickly unfolding neural responses that both precede and comprise the P2. The scalp-recorded ERP waveform represents the summation of neural activity that overlaps in time and space (Luck, 2005). PCA allows decomposition of this waveform into unique clusters of variance that meaningfully reflect distinct, underlying psychological processes (Dien and Frishkoff, 2005). Based on the DI model, we hypothesized that fixation, primarily representing differences in the influence of bottom-up processing, would have a large effect on early components but then diminish in subsequent components. Conversely, we hypothesized that race, which operationalizes more top-down differences in categorical perception, would have a small effect initially but then increase as neurocognitive iterations progressed.

Here, a multilevel modeling (MLM) approach was used to statistically test the effect of race and fixation on early-latency neural responses to faces. MLM has been advocated as more appropriate than repeated-measures ANOVA for psychophysiological data (Kristjansson *et al.*, 2007; Vossen *et al.*, 2011; Tibon and Levy, 2015; Tremblay and Newman, 2015), because (1) MLMs have more relaxed assumptions regarding sphericity, which psychophysiological data often violate; (2) MLMs allow

simultaneous parsing of variance associated with different grouping variables, including subjects, electrodes or stimulus items, thereby reducing error variance; (3) MLMs handle unbalanced or missing data, such that individuals with missing observations can be retained in the analysis; and (4) MLMs model effects of both categorical and continuous predictors simultaneously. These advantages make MLM a highly flexible and powerful analytic technique for ERP data (Page-Gould, 2017).

Traditionally, ERP responses are averaged over tens or hundreds of trials to extract the signal of interest (e.g. amplitude of a given component) from background EEG responses unrelated to stimulus processing (Luck, 2005). Given MLM's ability to handle unbalanced data and parse variance in a way that reduces error variance, data from individual trials can be modeled separately, thereby permitting examination of changes in the effects of interest over the course of many trials. While not directly pertinent to testing DI model predictions (which focus on events within trials), we present exploratory findings using this across-trials approach as a way of investigating stability and change in P2 amplitude in response to our manipulations over the course of the experiment.

Finally, the current study employed two different tasks to examine whether the task-relevance of person construal affects the applicability of the DI model. The first task was based on traditional evaluative priming paradigms (Fazio *et al.*, 1995; Livingston and Brewer, 2002), in which faces are irrelevant to the task of categorizing words as positive or negative. In contrast, faces were directly task-relevant in the second task as participants were asked to simply categorize them by race.

Methods

Participants

Sixty-five individuals (34 women, 31 men) participated in exchange for credit towards a research requirement in an Introductory Psychology course, or for monetary compensation. Participants ranged from 18 to 48 years old ($M = 20.4$). Sixty self-identified as White, two identified as Asian and three identified as more than one race. None identified as African-American.

Measures and procedure

Two computer tasks were administered using E-Prime (Psychology Software Tools, Inc., USA). Participants were seated ~40 inches from a 20-inch CRT monitor refreshing at 60 Hz. EEG data were recorded while each participant first completed the evaluative priming task and then the race categorization task.[1]

Evaluative priming task. The evaluative priming task was modified from tasks used previously (Fazio *et al.*, 1995) and is designed to measure bias in evaluative associations with African-American and European-American men. During each trial, a fixation cross was presented in the center of the screen (jittered: either 500, 700 or 900 ms), followed by a face prime (310 ms), then a blank screen (50 ms) and then a target word (200 ms), followed by a visual mask (600 ms). Prime stimuli consisted of photographs of Black and White men's faces with neutral expressions (taken from Ma *et al.*, 2015). In order to reduce differences in low-level perceptual features across faces, the photographs were converted to gray scale and the brightness and contrast of the images were adjusted to be roughly equivalent across stimuli; differences could not be completely

eliminated, however. Additionally, the location of the face prime varied so that the fixation cross preceded either the middle of the forehead or between the eyes (each face stimulus was presented once in each fixation position). Target stimuli consisted of positive and negative words that were somewhat visually degraded (see Supplementary Material for a complete list). Participants identified the valence of the target word using two keys on a ms-accurate keyboard using the index fingers of each hand; response mapping varied randomly across participants. Failure to respond within 800 ms of target onset elicited a 'TOO SLOW' warning displayed for 1000 ms. The ITI was 600 ms.

Participants completed 16 practice trials, followed by 512 experimental trials. Trial type (e.g. Black-eyes-positive word, Black-eyes-negative word, etc.) varied randomly, with 64 trials of each type in total. The same eight positive and eight negative words were used in the practice and experimental trials. Thirty-two faces of each race were used in the experimental trials; a different set of faces was used in the practice trials.

Race categorization task. In the race categorization task, participants viewed the same faces as in the experimental trials of the priming task, again presented in both fixation positions. Participants were asked to simply categorize the faces by race using two buttons on a keyboard. During each trial, a fixation cross was presented (jittered: 500, 700 or 900 ms), followed by a face (270 ms) presented either in the eyes-fixation or forehead-fixation position, which was then masked (530 ms). Failure to respond within 800 ms following target face onset elicited a 'TOO SLOW' warning displayed for 1000 ms. The ITI was 600 ms. Participants completed eight practice trials followed by 256 experimental trials. Trial type varied randomly, with 64 trials of each type being presented total.

Electrophysiological recording and processing

EEG data were collected using 20 tin electrodes embedded in a stretch-lycra cap (Electro-Cap, International, Eaton, OH) and placed in standard 10–20 locations (American Encephalographic Society, 1994).[2] All scalp electrodes were referenced online to the right mastoid; an average mastoid reference was derived offline. Signals were amplified with a Neuroscan Synamps amplifier (Compumedics, Charlotte, NC), filtered on-line at 10–40 Hz at a sampling rate of 1000 Hz. Impedances were kept below 10 KΩ. Ocular artifacts (i.e. blinks) were corrected from the EEG signal using a regression-based procedure (Semlitsch *et al.*, 1986). Trials containing voltage deflections of ±75 microvolts (μV) were discarded, as were trials that contained large muscle artifacts as determined by visual inspection.

P2 quantification. Grand averages (ERP activity averaged across trials and participants) revealed a positive-going deflection peaking roughly 160 ms following the presentation of a face and maximal at the centro-parietal midline (CPz), consistent with previous characterizations of the P2 during face processing (Ito and Urland, 2005; Dickter and Bartholow, 2007). The P2 was quantified in both tasks as the mean amplitude from 130 to 190 ms post-face onset (30 ms before and after the peak at CPz) at seven central and centro-parietal locations (Cz, C3, C4, CPz, CP3, CP4 and Pz).

Statistical approach. The R package 'lme4' (Bates *et al.*, 2015b) was used to fit multilevel models for data analysis. We allowed for covariances between random slopes and intercepts, using model-specification procedures described by Bates *et al.*, (2015a)

---

1 Participants also completed the Internal and External Motivation to be Unprejudiced Scale (Plant and Devine, 1998); the questionnaire data were not analyzed and will not be discussed.

2 Additional electrodes were placed above and below the left eye and on the outer canthus of each eye (to record blinks and saccades), over each mastoid, and on the tip of the nose.

to determine the most appropriate random effects structure. This involved starting with a maximal model and then removing random slopes based on the magnitude of the correlations between random effects. Estimated random effect variances and correlations can be found in the Supplementary Material. Satterthwaite approximations were used to estimate degrees of freedom and to obtain two-tailed *P* values; in situations where the degrees of freedom were above 200, we report the results as z statistics. Data and code used for analysis can be found at https://github.com/hiv8r3/ERP-fix-analyses.

Results

Only trials on which correct responses were given were used in analyses. Reaction time (RT) and ERP data from the priming task

Table 1. Mean reaction times (and SDs) and accuracy rates (and SDs) as a function of prime, target and fixation in each task separately

	Black primes	White primes
Evaluative priming task		
Target	Eyes fixation	
Positive word	502 (84)//0.91 (0.06)	505 (87)//0.91 (0.07)
Negative word	519 (85)//0.92 (0.07)	517 (82)//0.91 (0.07)
	Forehead fixation	
Positive word	502 (86)//0.92 (0.06)	503 (86)//0.91 (0.07)
Negative word	516 (83)//0.92 (0.06)	515 (82)//0.91 (0.08)
	Black targets	**White targets**
Race categorization task		
Fixation		
Eyes	451 (88)//0.93 (0.05)	456 (92)//0.93 (0.05)
Forehead	456 (92)//0.93 (0.06)	461 (93)//0.93 (0.05)

Note. Numbers in parentheses are standard deviations. Numbers to the left of forward slashes are mean (and SD) reaction times in milliseconds (correct response trials only). Numbers to the right of the slashes are mean accuracy rates.

for two subjects were discarded because accuracy was > 3 SDs below the mean (65.6% and 50.2%, respectively). Data from the categorization task for one subject were similarly discarded (60.9% accurate). Mean RTs and accuracy rates are presented in Table 1.

Reaction time

Evaluative priming task. Race of the face prime, valence of the target word and fixation were included in the model as predictors (dummy-coded: Black = 0, White = 1; negative = 0, positive = 1; eyes = 0, forehead = 1). The most appropriate random effects structure was determined to be one in which the intercept and effect (slope) of word valence varied by subject, and the intercept varied by stimulus. The Race × Word Valence interaction was significant, $b = 5.86$, $z = 2.37$, $P = 0.018$. The pattern of means associated with this interaction indicated that responses were faster to positive than negative words following both Black and White faces (Figure 1), but this facilitation effect was slightly (but significantly) larger following Black faces ($M = 15.5$ ms) compared to White faces ($M = 12$ ms). A main effect of Fixation also emerged, $b = -3.89$, $z = -2.23$, $P = 0.026$, such that words were evaluated more quickly following a forehead fixation than an eyes fixation. No other effects were significant; additional analyses can be found in the Supplementary Material.

Race categorization task. Race of the face prime and fixation were included as predictors (dummy-coded as before). The most appropriate random effects structure was determined to be one in which intercept and slopes of race and fixation (but not their interaction) varied by subject, and the intercept varied by face stimulus. A main effect of Fixation, $b = 5.78$, $z = 2.91$, $P = 0.004$, a marginal effect of Race, $b = 5.68$, $t(124) = 1.80$, $P = 0.074$ and no interaction, $b = -0.38$, $z = -0.144$, $P = 0.886$, emerged (Figure 1).

Primary ERP results: effects within trials

Traditional P2 amplitude analysis. We first tested the effects of race (Black = 0, White = 1) and fixation (eyes = 0, forehead = 1)
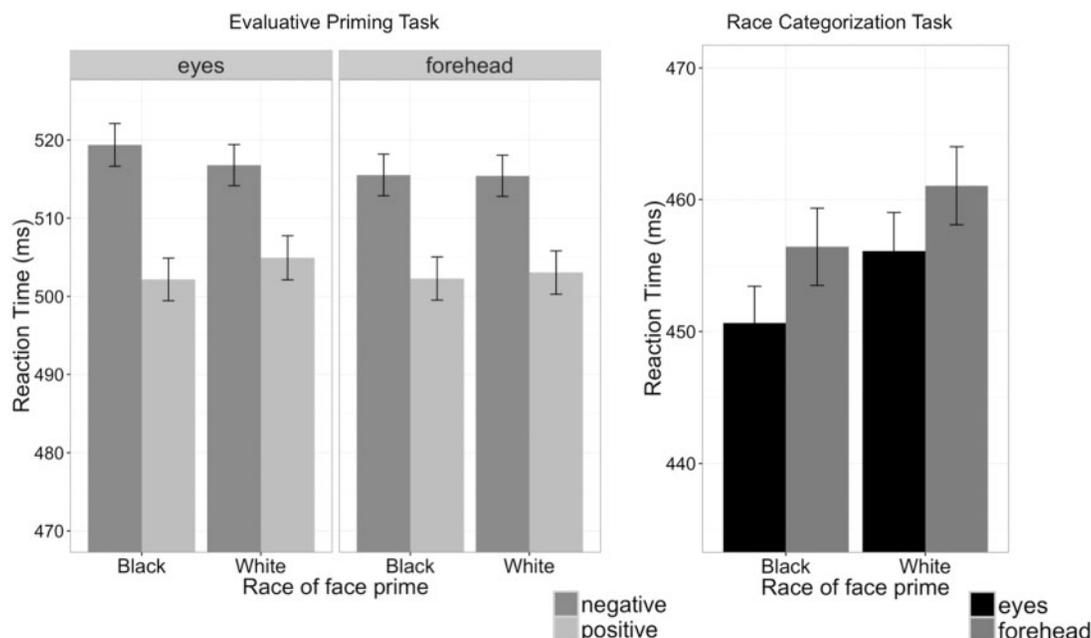

Fig. 1. Mean reaction times (correct responses) to target faces in each task as a function of condition. Error bars represent the standard error of the mean.
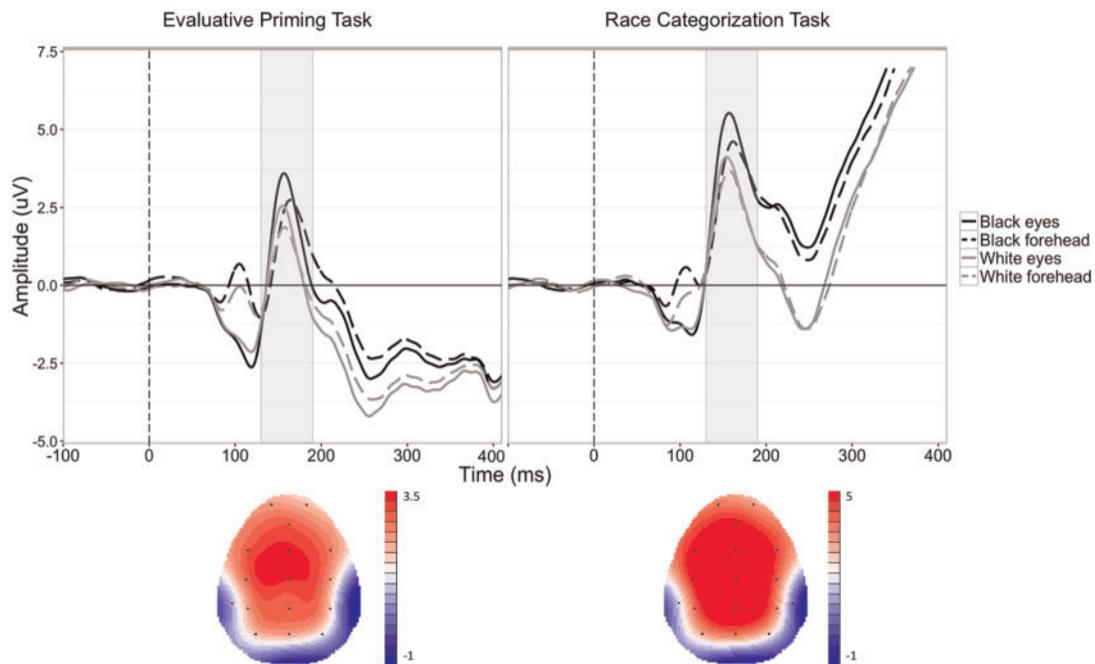
Fig. 2. Displays grand average waveforms locked to face onset for each task, averaged across C3, Cz, C4, CP3, CPz, CP4 and Pz. Positive amplitude is plotted upward. P2 mean amplitude was calculated from 130 to 160 ms following face presentation (shaded area).

on P2 amplitude using a traditional mean amplitude approach. Grand average ERP waveforms depicting the P2 are given in Figure 2. The random effects structure allowed the intercept, slopes of race, fixation and their interaction to vary by subject and the intercept to vary by electrode nested within subject. A significant main effect of Race was estimated in both the priming task, $b = -0.79$, $t(61.62) = -5.83$, $P < 0.001$, and the categorization task, $b = -1.20$, $t(63.36) = -4.81$, $P < 0.001$, such that Black faces elicited larger (more positive) P2s than White faces. A significant main effect of Fixation also emerged in both the priming task, $b = -0.39$, $t(62.15) = -2.03$, $P = 0.047$, and the categorization task, $b = -0.69$, $t(63.48) = -2.95$, $P = 0.005$; larger P2s were elicited in the eyes-fixation than the forehead-fixation condition. The Race $\times$ Fixation interaction was not significant in either task, $Ps > 0.34$. Additional analyses can be found in the Supplementary Material.

Principal components analysis. The primary hypothesis of the DI model (i.e. that the influence of variables representing bottom-up and top-down contributions changes as person construal progresses within individual trials) was tested by subjecting ERP responses to a sequential temporospatial PCA (Dien and Frishkoff, 2005), using the Matlab PCA ERP Toolbox (Dien, 2010). Separate PCAs were computed for the categorization and priming task data. Given that the presentation of the face was interrupted in the evaluative priming task after 360 ms, and because we were interested only in early person construal processes, we examined PCA components that emerged within 300 ms of face presentation. Details concerning extraction of components can be found in the Supplementary Materials. To facilitate interpretation of the PCA results, the portion of the original data set represented by each temporospatial factor combination was reconstructed (i.e. in microvolts) into factor waveforms by multiplying factor scores by their corresponding loadings and SDs. These reconstructed factor waveforms were then ordered temporally (henceforth referred to as Virtual Factors [VFs] 1 through 3

representing their temporal order) and viewed in comparison with the grand average ERPs (Figure 3).

To investigate the effects of race and fixation on each virtual factor, the mean amplitude of each factor was calculated separately for each condition and individual within the two tasks. In the evaluative priming task, VF-1, which peaked at 115 ms post-stimulus onset and was maximal at Pz, was quantified as mean amplitude 80–140 ms post-stimulus. VF-2, which peaked at 148 ms and was maximal at FCz, was quantified as mean amplitude 115–180 ms post-stimulus. VF-3 peaked at 179 ms and was maximal at CPz, and was quantified as mean amplitude 145–230 ms post-stimulus.[3] Mean VF amplitudes were subjected to MLMs with Race and Fixation (but not their interaction) as predictors and a random effects structure where the intercept and slopes of both effects varied by subject and the intercept varied by electrodes nested within subject. Predictors were effect-coded. Results across the three models revealed an increase in the (absolute-value) effect of race across the three virtual factors, while the (absolute value) effect of fixation decreased across the three virtual factors (Table 2, Figure 4). Specifically, the 95% confidence intervals for each estimate indicate a similar magnitude of the effect of Race on VF-1 and VF-2 but a statistical increase in the magnitude of the effect of Race from VF-2 to VF-3. In contrast, the magnitude of the effect of Fixation decreases from VF-1 to VF-3, although the magnitude of the effect on VF-2 does not statistically differ from either VF-1 or VF-3.

Using data from the race categorization task, a temporospatial PCA revealed three components that matched VF-1, VF-2 and VF-3 from the priming task in timing and location: VF-1 peaked at 113 ms post-stimulus and was maximal at Pz; VF-2 peaked at 143 ms and was maximal at FCz; and VF-3 peaked at 172 ms and was maximal at Cz. Because of these similarities and the fact that

---

3 VF-1 was quantified at Pz, CPz, Cz, P3, P4, CP3, CP4 and C3. VF-2 was quantified at FCz, Cz, Fz, CPz, FC4 and FC3. VF-3 was quantified at Cz, FCz, CPz, C3, FC3, Fz, Pz, CPz, FC4 and C4.
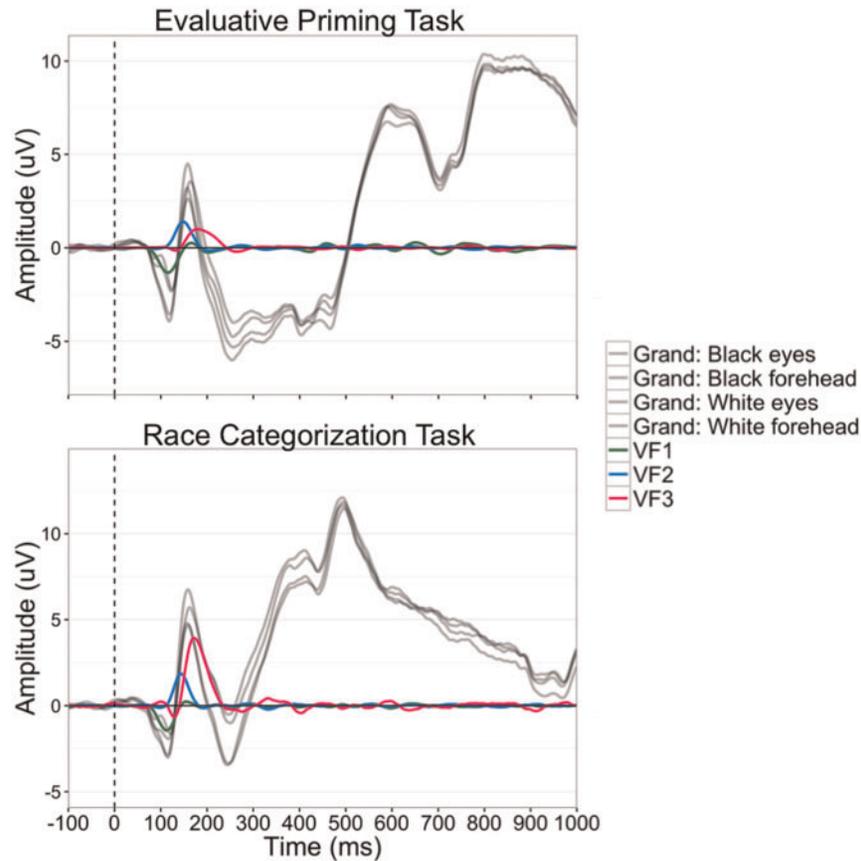
Fig. 3. Grand average waveforms elicited during both tasks overlaid with PCA-derived waveforms depicting virtual factor 1 through 3 in temporal order collapsed across conditions. The dashed vertical line indicates onset of the face.

they were elicited by the same face stimuli, these components were judged to represent similar processes across tasks. Quantification and analyses mirrored those for the priming task data, and a similar pattern was found: the effect of race increased as processing continued, while the effect of fixation decreased (Table 2, Figure 4). Examination of the 95% confidence intervals revealed the same pattern of results as in the categorization task.

Exploratory ERP results: effects across trials

Mean P2 amplitudes (130–190 ms post-face onset) from individual trials over the course of each task as a function of the race and fixation manipulations are plotted in Figure 5. Across both tasks, the data suggest an overall sensitization of the P2 (increasing across trials) and differing effects of race and fixation. Specifically, whereas the effect of race was evident from the earliest trials in both tasks, an effect of fixation emerged only as each task progressed such that P2s became larger in the eyes-fixation condition than the forehead-fixation condition. Moreover, race and fixation appeared to interact as the task progressed. These trends were confirmed by MLMs conducted separately with data from each task (Trial was added as a continuous predictor and rescaled to range from 0 to 10 in each task), the results of which are given in Table 3. The presence of significant Race × Fixation × Trial interactions in both models confirms that the slopes related to each effect differed, i.e. that the increases in P2 amplitude over the course of the tasks were asymmetrical across the four conditions. To probe this interaction, slope estimates and 95% confidence intervals were

calculated in accordance with Bauer and Curran (2005) (Table 4). All estimates are significantly different from zero, demonstrating positive change in P2 amplitude over the course of both tasks in all classes of stimuli. However, in both tasks, P2 amplitude in the Black-eyes condition increased more than in the other three conditions, as indicated by lack of overlap in the confidence intervals.

Discussion

The purpose of this study was to directly test elements of the DI model of person construal, using ERP data acquired while participants viewed faces of different races. The primary innovation of the DI model is its characterization of person construal as an iterative process in which bottom-up perceptual information is integrated with (top-down) stored representations related to social categories (Freeman and Ambady, 2011). A key assumption of this model is that bottom-up processes have a larger initial effect, while effects of top-down processes emerge later in processing. Here, this basic premise was tested using a fixation manipulation to control the visual information to which perceivers initially attended upon seeing faces of White and Black men.

We used multiple methods to investigate the ERP data from this study. A traditional mean amplitude approach to the P2 showed that, as in previous studies (Ito and Urland, 2003; Dickter and Bartholow, 2007) Black (outgroup) faces elicited larger P2s than White (ingroup) faces, regardless of fixation location. Additionally, fixating on the eyes elicited larger P2s

Table 2. Results of separate MLMs examining effects of race and fixation on mean amplitudes of PCA-derived virtual factors in both tasks

	VF-1	VF-2	VF-3
Evaluative priming task			
Race	0.08*	−0.09*	−0.39*
	[0.001, 0.15]	[−0.15, −0.03]	[−0.48, −0.30]
Fixation	0.53*	−0.34*	0.15*
	[0.42, 0.65]	[−0.46, −0.21]	[0.06, 0.24]
Race categorization task			
Race	−0.05	−0.10*	−0.46*
	[−0.15, 0.05]	[−0.19, −0.02]	[−0.57, −0.36]
Fixation	0.41*	−0.24*	−0.02
	[0.27, 0.54]	[−0.36, −0.13]	[−0.11, 0.08]

Note. Unstandardized betas are presented. Satterthwaite approximations were used to estimate degrees of freedom to calculate *P* values. Race and Fixation were both effect coded. Numbers in brackets are the lower and upper bounds of the 95% confidence interval around the estimate.
*$P < 0.05$.

emergence of VF-1 and its largely posterior scalp distribution suggest this component reflects activity in visual cortical circuits that is responsive to low-level stimulus features, such as the more complex spatial frequencies around the eyes relative to the forehead (Keil, 2009), and that is responsible for amplification of sensory information flowing to other parts of the visual attention pathway (Hillyard and Anllo-Vento, 1998). The temporal and spatial overlap between VF-3 and the P2 evident in the grand averages suggests that VF-3 directly contributed to the P2. This possibility is bolstered by the fact that the P2 is known to be highly sensitive to distinguishing social categories (Ito and Urland, 2003; Dickter and Bartholow, 2007), and that social category information (in this case, race) had a pronounced effect on VF-3 but a smaller effect on the preceding components.

More importantly, the increasing effect of race across the PCA-derived factors suggests that learned racial categories accessed from higher-level memory percepts contribute to the active representation of the social category in a top-down man-
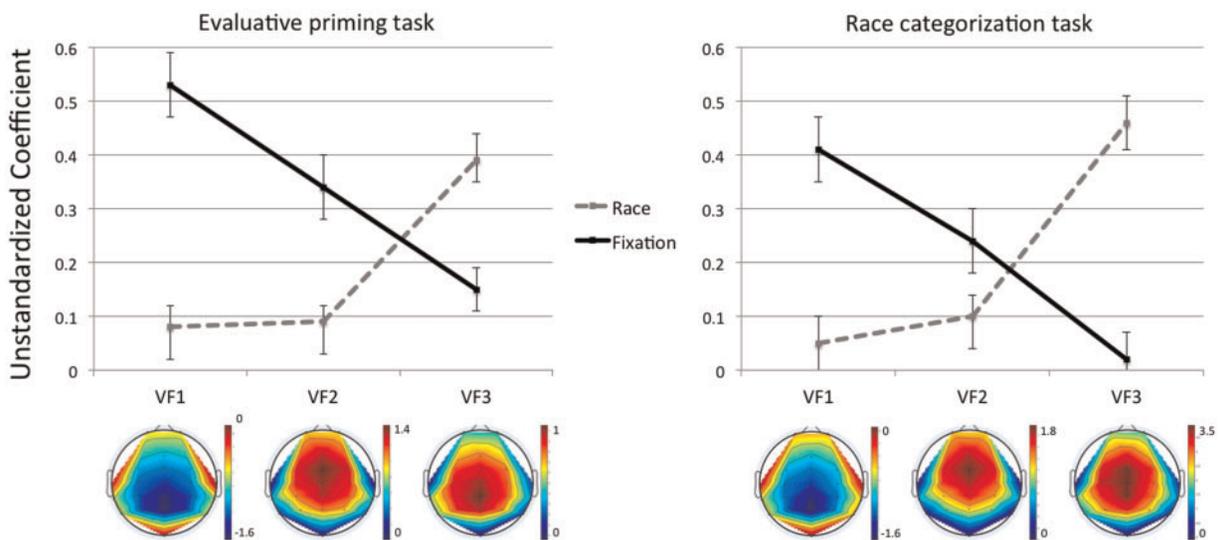

Fig. 4. Displays absolute values of unstandardized beta estimates for Race and Fixation effects from the three models predicting mean amplitude of PCA-derived Virtual Factors 1, 2 and 3 in each task. Error bars depict standard error of the estimate from the models. Corresponds to values in Table 2.

than fixating on the forehead. Although not predicted, this effect is consistent with evidence that faces with direct gazes are arousing and capture attention (Gale *et al.*, 1978; Senju and Hasegawa, 2005). Race and fixation location did not interact in this analysis, however.

Next, we examined the incorporation of bottom-up and top-down factors early in processing by testing the effects of race and fixation on a sequence of early components identified by a temporospatial PCA. In accordance with DI model predictions, we expected the effect of fixation (mainly representing differences in bottom-up processes) to be large upon initial perception of a face but to decrease over subsequent processing steps, whereas the effect of race (representing differences in a combination of top-down and bottom-up processes) was expected to be small initially but to increase over processing iterations. Consistent with these predictions, the effect of fixation was evident in the earliest component (80–140 ms following face onset) and decreased over the next 100 ms. In contrast, the effect of race on the first two components was small but increased dramatically in the third component. The very early

ner over time (Collins and Olson, 2014). Of course, it is important to acknowledge that stimulus features eliciting bottom-up and top-down processing were somewhat confounded in the current study. Given that race-related differences reflect both low-level, stimulus-driven and higher-level learned features, categorization by race represents a combination of bottom-up and top-down processes (Levin and Banaji, 2006). Indeed, the significant effect of race on the amplitude of VF-1 is likely due to low-level visual differences between faces of different races, despite efforts to equate stimuli on those dimensions. However, the increasing effect of race suggests learned racial categories accessed from higher-level memory percepts contribute to the active representation in an integrative way over time (Collins and Olson, 2014). Future research could extend this finding by using faces that do not differ in their low-level stimulus properties, as in a minimal groups design (Ratner and Amodio, 2013), to avoid bottom-up and top-down confounds. Another concern with the current design is that task order and the task-relevance of race categorization were confounded. Thus, inferences concerning the independence of the observed patterns
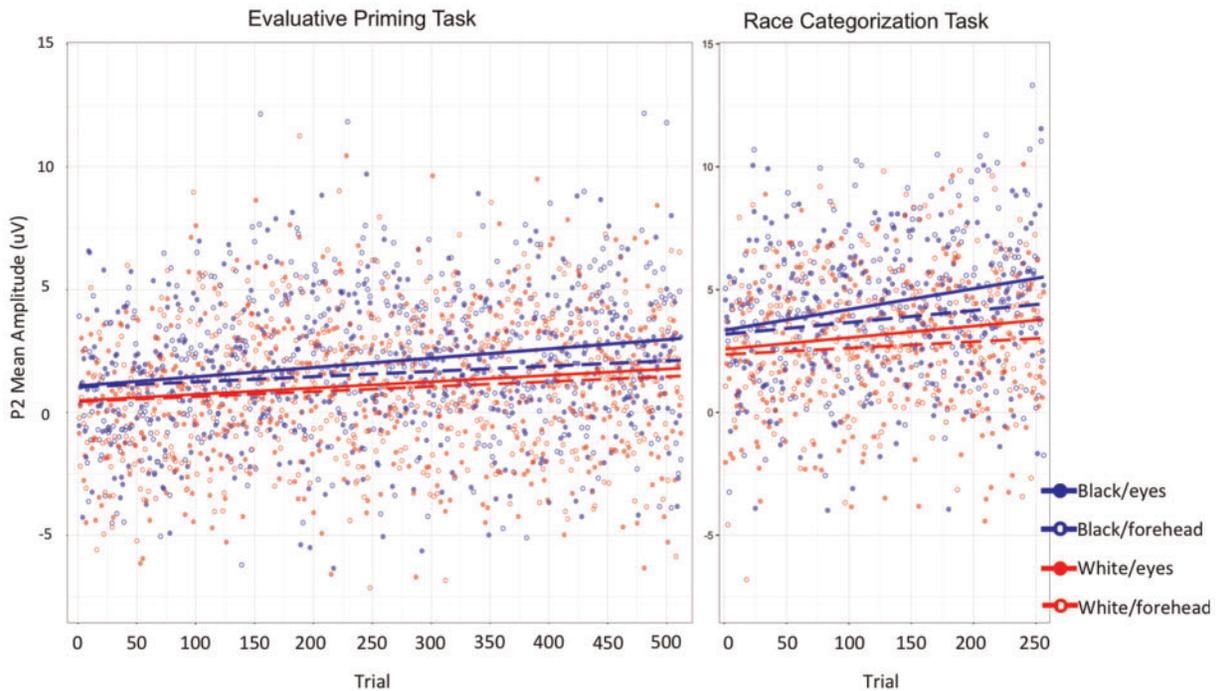
Fig. 5. P2 mean amplitude from individual trials (130–190 ms post-face onset) over the course of both tasks as a function of the race and fixation manipulations. For ease of viewing, each data point represents an average of all subjects for each trial for each condition separately. In the analysis, each data point for each subject was modeled separately.

Table 3. Results of multilevel models testing for change in the effect of race and fixation across trials on mean P2 amplitude

	Evaluative priming task		Race categorization task	
	b	*P*	*b*	*P*
Race	−0.46 (0.16)	**0.006**	−0.69 (0.28)	**0.015**
Fixation	−0.05 (0.22)	0.833	−0.004 (0.27)	0.998
Trial	0.17 (0.01)	**0.000**	0.21 (0.02)	**0.000**
Race × Fixation	−0.18 (0.23)	0.434	−0.20 (0.40)	0.618
Race × Trial	−0.07 (0.02)	**0.001**	−0.10 (0.03)	**0.000**
Fixation × Trial	−0.07 (0.02)	**0.000**	−0.14 (0.03)	**0.000**
Race × Fixation × Trial	0.07 (0.03)	**0.013**	0.11 (0.04)	**0.007**

Note. Unstandardized betas are presented. Standard errors of estimate are in parentheses. Satterthwaite approximations were used to estimate degrees of freedom to calculate *P* value. Variables were dummy coded; eyes = 0, forehead = 1; Black = 0, White = 1.

Table 4. Unstandardized coefficients and confidence intervals for the simple slope of trial on P2 amplitude as a function of condition in both tasks

	Black targets	White targets
Evaluative priming task		
Fixation		
Eyes	0.17 [0.14, 0.20]	0.11 [0.07, 0.13]
Forehead	0.10 [0.07, 0.13]	0.10 [0.08, 0.13]
Race categorization task		
Fixation		
Eyes	0.21 [0.17, 0.25]	0.11 [0.07, 0.15]
Forehead	0.07 [0.03, 0.11]	0.07 [0.03, 0.11]

Note. Numbers in brackets are the 95% confidence interval around the estimate. Trial has been rescaled to range between 0 and 10 for both tasks.

from perceivers' goals should be tempered. Still, the fact that such similar patterns emerged in both tasks is encouraging.

The exploratory analyses examining change in P2 amplitude across trials suggested that P2 amplitude increased across both tasks, and that this increase was most pronounced for Black-eyes trials. This pattern may be the result of participants' increasing ability to extract category-related information as the task continues, especially when focusing on the eyes—a highly practiced fixation location (Kawakami *et al.*, 2014)—and especially for outgroup faces, which elicit increased attention (Dickter and Bartholow, 2007). However, because inferences about the meaning of the face-elicited P2 are based on methods that assume the signal associated with the P2 is constant,

analyzing the P2 with this new approach impacts the (reverse) inferences we make about its psychological significance (e.g. Poldrack, 2006, but see Hutzler, 2014), and so we view this conclusion with caution. It is also important to emphasize that examination of change in P2 amplitude across trials is not relevant to testing DI model predictions, which focus on changes that should occur within trials (i.e. within construal events) as a function of various manipulations.

Interestingly, the P2's sensitivity to race was not reflected in priming task behavioral responses. The typical pattern of response facilitation for negative words following Black faces (Fazio *et al.*, 1995) was not seen in either fixation condition. Instead, participants were quicker to respond to positive than negative words in both race conditions, and this effect was consistent over the course of the task (see Supplementary Material). The phenomenon of evaluative priming is sensitive to a number

of parameters (Spruyt *et al.*, 2011). Thus, it could be that the SOA used here was too long to produce the behavioral priming phenomenon (see Supplementary Material for more extensive discussion). However, differentiation by race in the P2 and early PCA components provides evidence that categorization occurred, despite lack of behavioral evidence that this categorization had downstream consequences related to prejudice. These data are consistent with the idea that behavioral priming phenomena rely on response output processes, such as response conflict (Klinger *et al.*, 2000; Bartholow *et al.*, 2009), which are more sensitive to SOA and other task parameters (Spruyt *et al.*, 2007) than the initial categorization of faces.

In conclusion, the current study provides a novel demonstration using PCA that bottom-up and top-down processes integrate information in an iterative way to arrive at a stable person construal. Remarkably similar neural responses to faces were observed regardless of the relevance of social categorization for perceivers' task goals, suggesting automaticity of relevant construal processes. The temporal sensitivity of EEG and the ability for PCA to separate closely occurring but unique sources of variation in brain activity allows relatively direct access to this integration, which occurs before any behavioral response can be made, and speaks to the power of using covert measures of brain activity to investigate early and quickly unfolding processes of person construal.

Supplementary data

Supplementary data are available at *SCAN* online.

Funding

Acknowledgements

References

American Encephalographic Society (1994). Guideline thirteen: guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, **11**, 111–3.

Amodio, D.M. (2010). Coordinated roles of motivation and perception in the regulation of intergroup responses: frontal cortical asymmetry effects on the P2 event-related potential and behavior. *Journal of Cognitive Neuroscience*, **22**, 2609–17.

Amodio, D.M., Bartholow, B.D., Ito, T.A. (2014). Tracking the dynamics of the social brain: ERP approaches for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, **9**, 385–93.

Bartholow, B.D., Riordan, M.A., Saults, J.S., Lust, S.A. (2009). Psychophysiological evidence of response conflict and strategic control of responses in affective priming. *Journal of Experimental Social Psychology*, **45**, 655–66.

Bates, D., Kliegl, R., Vasishth, S., Baayen, R.H. (2015a) Parsimonious mixed models. arXiv:1506.04967.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015b). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48.

Bauer, D.J., Curran, P.J. (2005). Probing interactions in fixed and multilevel regression: inferential and graphical techniques. *Multivariate Behavioral Research*, **40**(3), 373–400.

Bodenhausen, G.V., Macrae, C.N. (1998). Stereotype activation and inhibition. In Wyer R. S. Jr., (Ed.), *Stereotype Activation and Inhibition: Advances in Social Cognition* (Vol. **11**, pp. 1–52). Hillsdale, NJ: Erlbaum.

Chen, J.M., Moons, W.G., Gaither, S.E., Hamilton, D.L., Sherman, J.W. (2014). Motivation to control prejudice predicts categorization of multiracials. *Personality and Social Psychology Bulletin*, **40**(5), 590–603.

Collins, J.A., Olson, I.R. (2014). Knowledge is power: how conceptual knowledge transforms visual cognition. *Psychonomic Bulletin & Review*, **21**(4), 843–60.

Correll, J., Urland, G.R., Ito, T.A. (2006). Event-related potentials and the decision to shoot: the role of threat perception and cognitive control. *Journal of Experimental Social Psychology*, **42**(1), 120–8.

Devine, P.G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, **56**, 5–18.

Dickter, C.L., Bartholow, B.D. (2007). Racial ingroup and outgroup attention biases revealed by event-related brain potentials. *Social Cognitive and Affective Neuroscience*, **2**(3), 189–98.

Dickter, C.L., Kittel, J.A. (2012). The effect of stereotypical primes on the neural processing of racially ambiguous faces. *Social Neuroscience*, **7**(6), 622–31.

Dien, J. (2010). The ERP PCA toolkit: an open source program for advanced statistical analysis of event-related potential data. *Journal of Neuroscience Methods*, **187**, 138–45.

Dien, J., Frishkoff, G.A. (2005). Introduction to principal components analysis of event-related potentials. In: Handy, T., editor. *Event Related Potentials: A Methods Handbook* (pp. 189–207). Cambridge, MA: MIT Press.

Di Russo, F., Martínez, A., Hillyard, S.A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex*, **13**(5), 486–99.

Dovidio, J.F., Kawakami, K., Johnson, C., Johnson, B., Howard, A. (1997). On the nature of prejudice: automatic and controlled components. *Journal of Experimental Social Psychology*, **33**, 510–40.

Fazio, R.H., Jackson, J.R., Dunton, B.C., Williams, C.J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, **69**(6), 1013–27.

Fiske, S.T., Neuberg, S.L. (1990). A continuum model of impression formation from category based to individuating processes: influences of information and motivation on attention and interpretation. In: Zanna, M.P., editor. *Advances in Experimental Social Psychology* (Vol. **3**, pp. 1–74). San Diego, CA: Academic Press.

Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Freeman, J.B., Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, **118**(2), 247–79.

Freeman, J.B., Ambady, N., Midgley, K.J., Holcomb, P.J. (2011). The real-time link between person perception and action: brain potential evidence for dynamic continuity. *Social Neuroscience*, **6**(2), 139–55.

Freeman, J.B., Penner, A.M., Saperstein, A., Scheutz, M., Ambady, N. (2011). Looking the part: social status cues shape race perception. *PLoS One*, **6**(9), 1–10.

Gale, A., Kingsley, E., Brookes, S., Smith, D. (1978). Cortical arousal and social intimacy in the human female under different conditions of eye contact. *Behavioural Processes*, **3**(3), 271–5.

Gilbert, C.D., Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, **14**(5), 350–63.

Hayward, W.G., Rhodes, G., Schwaninger, A. (2008). An own-race advantage for components as well as configurations in face recognition. *Cognition*, **106**(2), 1017–27.

He, Y., Johnson, M.K., Dovidio, J.F., McCarthy, G. (2009). The relation between race-related implicit associations and scalp-recorded neural activity evoked by faces from different races. *Social Neuroscience*, **4**(5), 426–42.

Hehman, E., Volpert, H.I., Simons, R.F. (2013). The N400 as an index of racial stereotype accessibility. *Social Cognitive and Affective Neuroscience*, **9**(4), 544–52.

Hills, P.J., Lewis, M.B. (2006). Reducing the own-race bias in face recognition by shifting attention. *The Quarterly Journal of Experimental Psychology*, **59**(6), 996–1002.

Hills, P.J., Lewis, M.B. (2011). Reducing the own-race bias in face recognition by attentional shift using fixation crosses preceding the lower half of a face. *Visual Cognition*, **19**(3), 313–39.

Hillyard, S.A., Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(3), 781–7.

Hutzler, F. (2014). Reverse inference is not a fallacy per se: cognitive processes can be inferred from functional imaging data. *Neuroimage*, **84**, 1061–9.

Ho, A.K., Sidanius, J., Cuddy, A.J., Banaji, M.R. (2013). Status boundary enforcement and the categorization of black–white biracials. *Journal of Experimental Social Psychology*, **49**(5), 940–3.

Ito, T.A., Bartholow, B.D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, **13**(12), 524–31.

Ito, T.A., Urland, G.R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, **85**(4), 616.

Ito, T.A., Urland, G.R. (2005). The influence of processing objectives on the perception of faces: an ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience*, **5**(1), 21–36.

Kawakami, K., Amodio, D.M., Hugenberg, K. (2017). Intergroup perception and cognition: an integrative framework for understanding the causes and consequences of social categorization. *Advances in Experimental Social Psychology*, **55**, 1–80.

Kawakami, K., Williams, A., Sidhu, D., *et al.* (2014). An eye for the I: preferential attention to the eyes of ingroup members. *Journal of Personality and Social Psychology*, **107**(1), 1–20.

Keil, M.S. (2009). "I look in your eyes, honey": internal face features induce spatial frequency preference for human face processing. *PLoS Computational Biology*, **5**(3), e1000329.

Klinger, M.R., Burton, P.C., Pitts, G.S. (2000). Mechanisms of unconscious priming I: Response competition not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **26**, 441–55.

Kristjansson, S.D., Kircher, J.C., Webb, A.K. (2007). Multilevel models for repeated measures research in psychophysiology: an introduction to growth curve modeling. *Psychophysiology*, **44**, 728–36.

Krosch, A.R., Amodio, D.M. (2014). Economic scarcity alters the perception of race. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(25), 9079–84.

Krosch, A.R., Berntsen, L., Amodio, D.M., Jost, J.T., Van Bavel, J.J. (2013). On the ideology of hypodescent: political conservatism predicts categorization of racially ambiguous faces as Black. *Journal of Experimental Social Psychology*, **49**(6), 1196–203.

Kubota, J.T., Ito, T.A. (2007). Multiple cues in social perception: the time course of processing race and facial expression. *Journal of Experimental Social Psychology*, **43**(5), 738–52.

Lepore, L., Brown, R. (1997). Category and stereotype activation: is prejudice inevitable? *Journal of Personality and Social Psychology*, **72**(2), 275–87.

Levin, D.T., Banaji, M.R. (2006). Distortions in the perceived lightness of faces: the role of race categories. *Journal of Experimental Psychology: General*, **135**(4), 501–12.

Livingston, R.W., Brewer, M.B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology*, **82**(1), 5–18.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: MIT Press.

Ma, D.S., Correll, J., Wittenbrink, B. (2015). The Chicago Face Database: A Free Stimulus Set of Faces and Norming Data. *Behavior Research Methods*, **47**, 1122–35.

Macrae, C.N., Milne, A.B., Bodenhausen, G.V. (1994). Stereotypes as energy-saving devices: a peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, **66**, 37–47.

Page-Gould, E. (2017). Multilevel modeling. In: Cacioppo, J.T., Tassinary, L., and Berntson, G., editors. *The Handbook of Psychophysiology*, 4th edn. New York: Cambridge University Press.

Peterson, M., Kanwisher, N. (2015). Where do people look on faces in the real world? *Journal of Vision*, **15**(12), 696.

Plant, E.A., Devine, P.G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, **75**(3), 811–32.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, **10**(2), 59–63.

Ratner, K.G., Amodio, D.M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology*, **49**, 298–301.

Rodeheffer, C.D., Hill, S.E., Lord, C.G. (2012). Does this recession make me look black? The effect of resource scarcity on the categorization of biracial faces. *Psychological Science*, **23**(12), 1476–8.

Semlitsch, H.V., Anderer, P., Schuster, P., Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, **23**(6), 695–703.

Senju, A., Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cognition*, **12**(1), 127–44.

Spruyt, A., Gast, A., Moors, A. (2011). The sequential priming paradigm: A primer. In Klauer, K.C., Voss, A., Stahl C., editors. *Cognitive Methods in Social Psychology*, pp. 48–77. New York, NY: Guilford Press.

Spruyt, A., Hermans, D., De Houwer, J., Vandromme, H., Eelen, P. (2007). On the nature of the affective priming effect: effects of stimulus onset asynchrony and congruence proportion in naming and evaluative categorization. *Memory and Cognition*, **35**, 95–106.

Stanley, D.A., Sokol-Hessner, P., Banaji, M.R., Phelps, E.A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(19), 7710–5.

Stolier, R.M., Freeman, J.B. (2016a). Neural pattern similarity reveals the inherent intersection of social categories. *Nature Neuroscience*, **19**(6), 795–7.

Stolier, R.M., Freeman, J.B. (2016b). Functional and temporal considerations for top-down influences in social perception. *Psychological Inquiry*, **27**(4), 352–7.

Teufel, C., Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. *Consciousness and Cognition*, **47**, 17–25.

Tibon, R., Levy, D.A. (2015). Striking a balance: analyzing unbalanced event-related potential data. *Frontiers in Psychology*, **6**, 555.

Tremblay, A., Newman, A.J. (2015). Modeling nonlinear relationships in ERP data using mixed-effects regression with R examples. *Psychophysiology*, **52**, 124–39.

Tskhay, K.O., Rule, N.O. (2015). Semantic information influences race categorization from faces. *Personality and Social Psychology Bulletin*, **41**, 769–78.

Vetter, P., Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Consciousness and Cognition*, **27**, 62–75.

Vossen, H., Van Breukelen, G., Hermans, H., Van Os, J., Lousberg, R. (2011). More potential in statistical analyses of event-related potentials: a mixed regression approach. *International Journal of Methods in Pediatric Research*, **20**(3), 56–68.

Willadsen-Jensen, E.C., Ito, T.A. (2006). Ambiguity and the time-course of racial perception. *Social Cognition*, **24**(5), 580–606.

Willadsen-Jensen, E.C., Ito, T.A. (2008). A foot in both worlds: Asian Americans' perceptions of Asian, White, and racially ambiguous faces. *Group Processes & Intergroup Relations*, **11**(2), 182–200.

Zhao, L., Bentin, S. (2011). The role of features and configural processing in face-race classification. *Vision Research*, **51**(23), 2462–70.