Journal of Experimental Psychology: General
2012, Vol. 141, No. 4, 715–727

© 2012 American Psychological Association
0096-3445/12/$12.00 DOI: 10.1037/a0027950

Unconscious Evaluation of Faces on Social Dimensions

Lorna H. Stewart, Sara Ajina, and Spas Getov
University College London

Bahador Bahrami
University College London and Aarhus University

Alexander Todorov
Princeton University

Geraint Rees
University College London

It has been proposed that two major axes, dominance and trustworthiness, characterize the social dimensions of face evaluation. Whether evaluation of faces on these social dimensions is restricted to conscious appraisal or happens at a preconscious level is unknown. Here we provide behavioral evidence that such preconscious evaluations exist and that they are likely to be interpretations arising from interactions between the face stimuli and observer-specific traits. Monocularly viewed faces that varied independently along two social dimensions of trust and dominance were rendered invisible by continuous flash suppression (CFS) when a flashing pattern was presented to the other eye. Participants pressed a button as soon as they saw the face emerge from suppression to indicate whether the previously hidden face was located slightly to the left or right of central fixation. Dominant and untrustworthy faces took significantly longer time to emerge (T2E) compared with neutral faces. A control experiment showed these findings could not reflect delayed motor responses to conscious faces. Finally, we showed that participants' self-reported propensity to trust was strongly predictive of untrust avoidance (i.e., difference in T2E for untrustworthy vs neutral faces) as well as dominance avoidance (i.e., difference in T2E for dominant vs neutral faces). Dominance avoidance was also correlated with submissive behavior. We suggest that such prolongation of suppression for threatening faces may result from a passive fear response, leading to slowed visual perception.

Keywords: continuous flash suppression, dominance, trust, social, preconscious

Supplemental materials: http://dx.doi.org/10.1037/a0027950.supp

Humans spontaneously and effortlessly evaluate the faces of other people whom they might have never known or met before on numerous social qualities, such as trust, competence, and friendliness. These inferences are extremely fast and quantitatively replicable across time (Todorov, Said, Engell, & Oosterhof, 2008). Moreover, these evaluations predict important social outcomes such as election results (Todorov, Mandisodza, Goren, & Hall, 2005). A recent study suggested that such evaluations may usefully predict the face owner's social behavior too: more trustworthy-looking men were more likely to reciprocate favors in a social economic game that called for mutual trust and cooperation (Stirrat & Perrett, 2010). It has been proposed that two major axes, trustworthiness and dominance, predominantly characterize the social dimensions of face evaluation (Oosterhof & Todorov, 2008).

Whether social evaluation of faces is restricted to conscious appraisal or happens at a preconscious level is unknown. Extensive behavioral and neurophysiological evidence suggests that processing of faces (as a category of visual object) is not restricted to conscious awareness (Jiang, Costello, & He, 2007; Sterzer & Rees, 2008; Sterzer, Jalkanen, & Rees, 2009). The case for preconscious face processing is perhaps strongest (and most popularized) for affective expressions such as fear (W. J. Adams, Gray, Garner, & Graf, 2010; Alpers & Gerdes, 2007; Jiang & He, 2006; Morris, Ohman, & Dolan,1998; Whalen et al., 2004; Yang, Zald, & Blake,

---

Lorna H. Stewart, Department of Clinical, Educational, and Health Psychology, University College London, London, England; Sara Ajina, and Spas Getov, Institute of Cognitive Neuroscience, University College London; Bahador Bahrami, Institute of Cognitive Neuroscience, University College London, Interacting Minds Project, Institute of Anthropology, Archaeology, and Linguistics, Aarhus University, Aarhus, Denmark, and Centre of Functionally Integrative Neuroscience, Aarhus University Hospital; Alexander Todorov, Department of Psychology, Princeton University; Geraint Rees, Institute of Cognitive Neuroscience, University College London, and Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London.

This work was supported by a British Academy postdoctoral fellowship (to Bahador Bahrami), the Danish National Research Foundation and the Danish Research Council for Culture and Communication, the European Union MindBridge project (award to Bahador Bahrami), the National Science Foundation of the United States (Grant NSF-0823749 to Alexander Todorov), a Medical Research Council studentship (to Lorna H. Stewart), a United Kingdom National Institute for Health Research academic clinical fellowship (to Spas Getov), and the Wellcome Trust (Geraint Rees). Support from the MIND*Lab* UNIK initiative at Aarhus University was funded by the Danish Ministry of Science, Technology, and Innovation. We thank Karsten Olsen for generous help with data collection in Experiments 2 and 3.

Correspondence concerning this article should be addressed to Bahador Bahrami, Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, United Kingdom. E-mail: bbahrami@gmail.com

2007; Yoon, Hong, Joormann, & Kang, 2009). The success of this line of research has led to proposition of a dedicated, automatic subcortical neuronal mechanism involving the human superior colliculus and amygdala for "quick-and-dirty" visual evaluation of evolutionarily relevant (negative) affective facial expressions (Morris, DeGelder, Weiskrantz, & Dolan, 2001; Morris, Ohman, & Dolan, 1999; Vuilleumier, 2005). This evolutionary vigilance account posits that preconscious face processing promotes survival by contributing to rapid, albeit coarse, detection of threatening situations. In addition to subcortical face processing, unconscious face processing can also occur in the cortex (e.g. Sterzer et al., 2009), and so there are several routes by which face processing could potentially influence behavior.

Two independent studies (Jiang, Costello, & He, 2007; Yang et al., 2007) recently introduced a novel method for assessing preconscious face processing. Jiang, Costello, and He (2007) presented an upright or inverted face to one eye of their participants and rendered the face invisible by simultaneously presenting rapidly flashing high-contrast masks (continuous flash suppression, or CFS) to the other eye. To obtain a measure of preconscious face processing, they measured the time it took the suppressed face to break through the suppression and become visible. Upright faces broke through CFS faster than inverted ones, indicating that preconscious face processing distinguishes between upright and inverted faces. Employing a very similar method, Yang and colleagues (2007) showed that fearful faces break through interocular suppression faster than happy and neutral faces, lending support to the vigilance hypothesis discussed previously. This paradigm provides a great opportunity to test predictions of the vigilance hypothesis about the preconscious social evaluation of faces. One prediction is that faces associated with stronger threat (i.e., more dominant and less trustworthy faces) should break through suppression faster than neutral faces.

In contrast, some studies in which face adaptation paradigms were used have cast doubt on the notion that higher level characteristics of faces can be coded unconsciously. Adaptation to facial identity (Moradi, Koch, & Shimojo, 2005), gender, and race (Amihai, Deouell, & Bentin, in press) are only observed when the adaptor faces are consciously perceived. Interestingly, invisible adaptor faces can induce adaptation to gender (Shin, Stolte, & Chong, 2009) and facial expression (Yang, Hong, & Blake, 2010) but only if spatial attention is directed to the location of the invisible adaptor. In the absence of any explicit spatial attentional cues, the results from the face adaptation literature (Amihai et al., in press; Moradi et al., 2005) suggest that higher level social evaluation of faces is restricted to conscious perception (but also see W. J. Adams et al., 2010, for a more recent report challenging this). In sum, the contradictory findings about preconscious high-level face perception show clearly that preconscious evaluation of faces on social dimensions is an open and highly relevant question.

In the present study, we asked whether evaluation of faces on social dimensions extended to preconscious perceptual processing and to what extent these preconscious processes were observer-specific. The two-dimensional trust-by-dominance face space (Oosterhof & Todorov, 2008) provides a well-controlled quantitatively validated stimulus repertoire of faces that allowed us to manipulate these two social dimensions independently (see Figure 1). To probe preconscious face processing, in Experiment 1, we rendered the faces invisible by interocular suppression and measured the time taken by the face to break into awareness (Jiang, Costello, and He, 2007; Yang et al., 2007). Experiment 2 was conducted as a control to confirm that the results of Experiment 1 could not be due to generalized difference in conscious responses to threatening faces. Finally, to assess observer-specific effects on preconscious perception, we employed three validated self-report questionnaires that assess propensity to show interpersonal trust (Evans & Revelle, 2008), frequency of submissive behaviours (Allan & Gilbert, 1997) and social anxiety (Leary, 1983). We tested whether these self-reported personality traits were related to individual variability in preconscious social evaluation.

Experiment 1

Method

Participants. Twenty-seven participants (11 women) took part in Experiment 1. All participants had normal or corrected-to-normal vision and ranged in age from 18 to 29 years (mean age 22.0 ± 3.1). Participants gave written informed consent according to the guidelines of the local research ethics committee. Four participants were excluded from the analysis. Two decided to leave the experiment in the middle. One had forgotten his glasses and could not see the faces clearly. Another participant revealed upon debriefing that she had resorted to blinking her CFS eye selectively to do the task.

Display apparatus and stimuli. We programmed the experimental paradigm using the Cogent Toolbox (Cogent 2000 Version 1.25; available from http://www.vislab.ucl.ac.uk/cogent.php/) for MATLAB Version 7.2.0.232 (R2006a; Mathworks, Natick, MA). Stimuli were presented on a Sony Trinitron GDM-F520 monitor (800 × 600 pixels at 85 Hz; Sony Corp., Tokyo, Japan) and viewed through a mirror stereoscope mounted on a head and chin rest, at a viewing distance of 65.5 cm. The images presented to both eyes of the participants were displayed side-by-side on the monitor, each with a tile frame surround (11.77° visual angle), upon a uniform gray background (background luminance = 65 Cd/m^2). A central white fixation cross (0.6° visual angle) was presented to each eye. Optimal fusion of the two images was ensured prior to commencing each experiment. Behavioral testing was carried out in a dark and quiet room.

Images of emotionally neutral White faces were generated with a customized version of the Facegen Modeller program (Figure 1A; available at http://facegen.com; Singular Inversions, Toronto, ON, Canada). This customized software version provides two orthogonal parameters that allowed us to manipulate perceived trustworthiness and dominance based on an extensively validated model (Oosterhof & Todorov, 2008). We systematically varied the trustworthiness and dominance of the same face identity in seven steps (−3,−2,−1, 0,+1,+2,+3) of 1 standard deviation straddling the neutral. The result was a set of 49 faces covering all possible combinations of trust and dominance in the employed range. Each face covered an ellipsoid area subtending 4.2° (height) by 2.6° (width).

Behavioral procedures. The experimental paradigm is shown in Figure 1B. Each trial started with presentation of CFS to the nondominant eye at full contrast. CFS consisted of dynamic noise patterns (frequency 10 Hz) generated by superimposition of shapes of random size and color at maximum contrast (Bahrami,

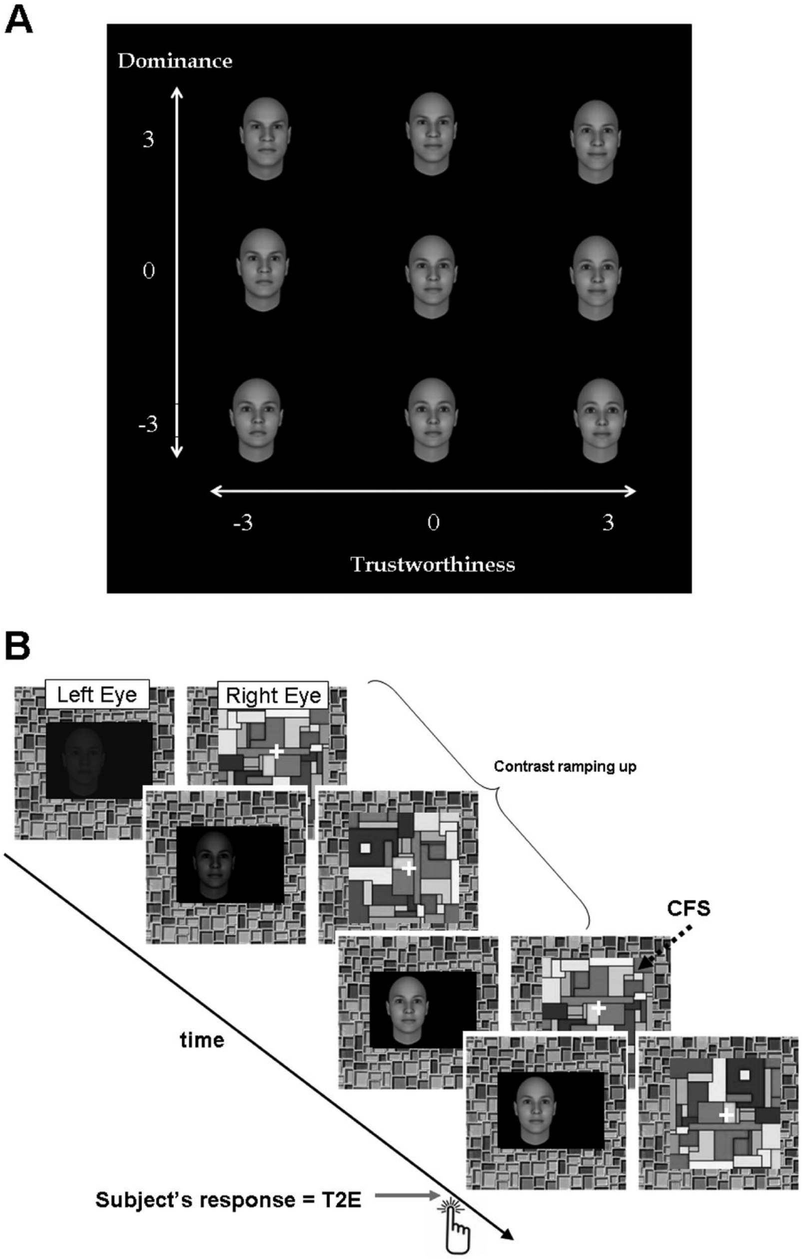

Figure 1. (A) Two-dimensional (Trust × Dominance) space of social evaluation of faces. (B) The sequence of events in a typical trial of Experiment 1. T2E = Time to emerge; CFS = continuous flash suppression.

Carmel, Walsh, Rees, & Lavie, 2008a, 2008b; Bahrami et al., 2010; Bahrami, Lavie, & Rees, 2007; Jiang, Costello, Fang, Huang, & He, 2006; Jiang, Costello, & He, 2007; Jiang, Zhou, & He, 2007; Tsuchiya & Koch, 2005). While CFS was presented to the nondominant eye, the test face was presented on a black background to the dominant eye at a location 1 cm (~0.7° visual angle) left or right of center. The contrast of the test face was ramped up linearly by a 4.5% increment every 100 ms from 0% to 100% during the initial 2,200 ms of the trial and subsequently remained constant (as per Jiang, Costello, & He, 2007) until the participant responded or the trial terminated. Participants were instructed to press a key(*C* or *Z*) on a standard keyboard as soon as they were confident that the suppressed face was either on the left or on the right side of the central fixation. Speed and accuracy were both emphasized. Thus, correct responses provided a measure of the time (from onset of stimulus presentation to key press) taken by the suppressed stimulus to break through the CFS and emerge into awareness. This measure, which we call *time to emerge* (*T2E*), was our main dependent variable for the face stimulus. If the face did not break through CFS or the participant did not respond after 10 s, the trial terminated automatically and was excluded from analysis. After every correct response, participants were asked to use the mouse pointer to click on the part of the face that had emerged through CFS first. The data from this task were corrupted by a programming error and are not reported here.

Participants completed a total of 490 trials (10 blocks of 49 trials each), with each of the 49 face stimuli presented a total of 10 times (once in each block). Each participant started with a practice block of 49 trials identical to the rest of the experiment, except that the eye presented with the face stimulus was randomized for each trial. This practice served the dual purpose of familiarizing participants with the paradigm and defining eye dominance. In a pilot test prior to the experiment, we had observed that when the suppressed face was presented to the nondominant eye, very occasionally the face did not break through the CFS even after 10 s. Thus presenting the suppressed face to the nondominant eye or indeed randomizing the eye to which the presentation of the suppressed face was made led to considerable variability across trials, and many trials had to be excluded. Using the practice block to determine the dominant eye gave us a consistent criterion for avoiding this undesirable variability. We defined the dominant eye as the eye for which the suppressed face broke through the CFS and emerged to consciousness more quickly. Throughout the main experiment, the suppressed stimulus was only presented to the dominant eye.

Data preparation. Only data from correct trials were analyzed. Error trials were rare. Mean accuracy was high (95.2%), and all participants achieved greater than 85% accuracy. A 2 (participant gender) × 7 (trust levels) × 7 (dominance levels) analysis of variance (ANOVA) showed no main effect of gender, $F(1, 21) = 0.92$, $p = .35$, and no interactions between gender and the other factors: trust: $F(6, 126) = 0.76$, $p = .60$, and dominance: $F(6, 126) = 1.21$, $p = .31$. Consequently, data were collapsed across gender for the remainder of analyses.

Results and Discussion

A two-way repeated measure ANOVA with dominance and trust as factors (seven levels for each factor) revealed a highly significant main effect of dominance, $F(6, 132) = 4.26$, $p = .001$, partial eta squared = .162, Figure 2A, and a significant main effect of trust, $F(6, 132) = 3.36$, $p = .004$, partial eta squared = .133; Figure 2B. The interaction between the two factors was not significant, $F(36, 792) = 0.85$, partial eta squared = .03; $p > .5$. As can be appreciated from the graphs in Figure 2, for dominance, the variance consisted of a significant linear—$F(1, 22) = 21.71$, $p < .001$, partial eta squared = .49—as well as a significant quadratic

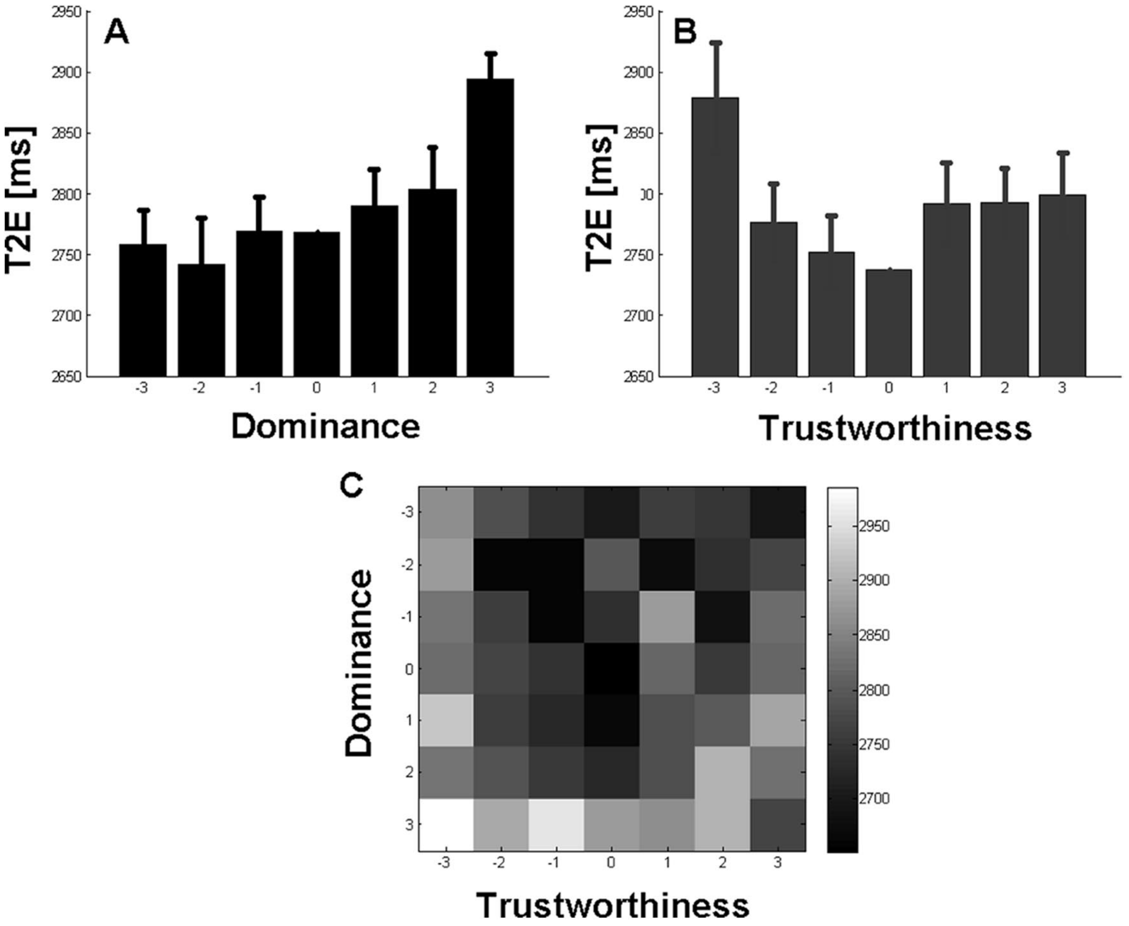

Figure 2. Experiment 1 results. Time to emerge (T2E) averaged across subjects ($N = 23$) are plotted against dominance (A) and trustworthiness (B) of the suppressed faces. Error bars = standard error of mean difference between the specific condition and neutral face. (C) Gray scale heat map illustrates the T2E for all 7 × 7 combinations of trust by dominance. Brighter shades correspond to larger T2E.

component—$F(1, 22) = 8.78$, $p = .007$, partial eta squared = .28. For trust, however, only the quadratic component was significant, $F(1, 22) = 6.82$, $p = .016$, partial eta squared = .237. For a detailed table of the mean raw T2E and standard deviations for each of the 7 × 7 combinations of trust and dominance, see Table 1 in the online supplementary material.

Planned pairwise comparisons revealed that the main effect of dominance in Experiment 1 reflected significantly longer T2E for the most dominant, $t(22) = 6.13$, $p < 10^{-4}$, Cohen's $d = 2.61$, effect size correlation $r = .79$, compared with the neutral–dominance faces. Furthermore, the least (-3 *SD*s) trustworthy faces took significantly longer time to break through consciousness compared with neutral faces—planned pairwise comparison of untrustworthy and neutral–trustworthy faces: $t(22) = 3.04$, $p < .007$; Cohen's $d = 1.29$, effect size correlation $r = .54$)—and there was a trend for most trustworthy faces to break through CFS later than neutral faces, $t(22) = 1.77$; $p = .09$; Cohen's $d = 0.75$, effect size correlation $r = .35$.

These results clearly showed that evaluation of faces on social dimensions extends to preconscious perceptual processing. More important, they provide evidence against the predictions of the vigilance hypothesis by showing that the most threatening faces (i.e., the most dominant and least trustworthy faces) broke through interocular suppression significantly more slowly than the neutral faces. Another unexpected finding was that the most trustworthy faces tended to break through CFS more slowly than neutral faces. As we will see further on, although this result was not predicted in the context of vigilance theory, it turns out to be robust and replicable across studies. We will discuss some possible interpretations of this finding in the General Discussion.

Experiment 2

The results of Experiment 1 raise an important alternative hypothesis: the observed results may not show a difference in duration of suppression under CFS but rather indicate that the observers may have been generally slower to respond to more threatening faces *after* the face had broken through CFS. This view would predict that a similar behavioral result would be observed even if the faces were presented consciously (i.e., without being suppressed behind CFS). We performed Experiment 2 to test this general "threat-induced slowing of motor response" hypothesis and replicate our main findings in a new group of participants. We compared the response times for correct discrimination of face locations when they were suppressed from awareness (monocular presentation of the face with CFS in the other eye—replication of Experiment 1) with when the faces were never suppressed (binocular presentation of the faces on top of CFS).

Method

Participants. Twenty-one participants (16 women) took part in Experiment 2. All participants had normal or corrected-to-normal vision and ranged in age from 18 to 35 years (mean age 22.4 ± 4.4). Participants gave written informed consent, and the experiment was approved by the local research ethics committee.

Display apparatus and stimuli. Display properties were identical to Experiment 1. For the face stimuli, given that in Experiment 1 the most prominent effects were found in the extremes of trust and dominance scales, here, manipulations of trustworthiness and dominance were restricted to (-3, 0, $+3$) standard deviations on each axis, producing a set of nine faces, covering the extremes as well as the middle point of each axis. Contrast was ramped up in the same manner as in Experiment 1.

Behavioral procedure. One half of the trials (randomly interleaved) were identical to Experiment 1. In the other half, the same face stimuli were presented to *both* eyes and *on top* of the CFS stimulus, leading to conscious, unsuppressed perception of the faces. The experiment consisted of 15 blocks of 18 trials. Each face stimulus was repeated twice in each block (once suppressed by CFS and once not). The participant's task was to decide whether the face was on the left or right side of the fixation point by pressing the same keyboard buttons as in the previous experiment.

Results and Discussion

Errors were rare. Mean accuracy was 99% in the visible condition and 96% in the suppressed condition. Analysis of response times was restricted to correct trials. The experimental manipulations consisted of two social (three levels of dominance and three levels of trust) and one nonsocial (two levels of visibility) factors (see Figure 3). A three-way (Dominance × Trust × Visibility) repeated-measure ANOVA was applied to the normalized correct T2E. A highly significant main effect of visibility was found, $F(1, 20) = 106.87$, $p < 0.001$; partial eta squared = .84, reflecting the trivial fact that response times were much faster to unsuppressed faces. More important, the main effect of dominance (Figure 3A) was significant, $F(1, 20) = 4.20$, $p = .02$; partial eta squared = .174. A significant interaction, $F(2, 40) = 3.38$, $p = .04$; partial eta squared = .14, was also observed between dominance and CFS, indicating that only in the suppressed condition did the highly dominating face take longer time to break through CFS. Post hoc comparison showed that T2E was significantly longer, $t(20) = 3.45$, $p = .003$; Cohen's $d = 1.54$; effect size correlation $r = .61$, for dominant than for neutral invisible faces, thus replicating the results of Experiment 1. The main effect of trust was also significant, $F(2, 40) = 6.69$, $p = .003$; partial eta squared = .25. The interaction between trust and suppression was significant, $F(2, 40) = 4.76$, $p = .01$, partial eta squared = .192. A significant quadratic component was observed for trust, $F(1, 20) = 11.74$; $p = .003$; partial eta squared = .370, which interacted significantly with visibility, $F(1, 20) = 7.54$; $p = .012$; partial eta squared = .274. Replicating Experiment 1, T2E was significantly longer both for untrustworthy faces, $t(20) = 2.99$, $p = .007$; Cohen's $d = 1.33$, effect size correlation $r = .55$, and trustworthy faces, $t(20) = 2.69$, $p = .01$; Cohen's $d = 1.2$, effect size correlation $r = .51$, compared with the neutral invisible faces. Neither the two-way interaction between trust and dominance ($p > .36$) nor the three-way interaction ($p > .17$) was significant. The detailed table of the mean raw correct T2E and standard deviations for each of the 3x3 combinations of trust and dominance for suppressed and visible faces are reported in Table 2 in the online supplementary material.

These results replicated the findings of Experiment 1 in the CFS condition in a new group of observers and showed that under conscious viewing conditions, reaction times were not modulated by social valence of the faces. The positive evidence from the significant interactions between dominance and visibility and a

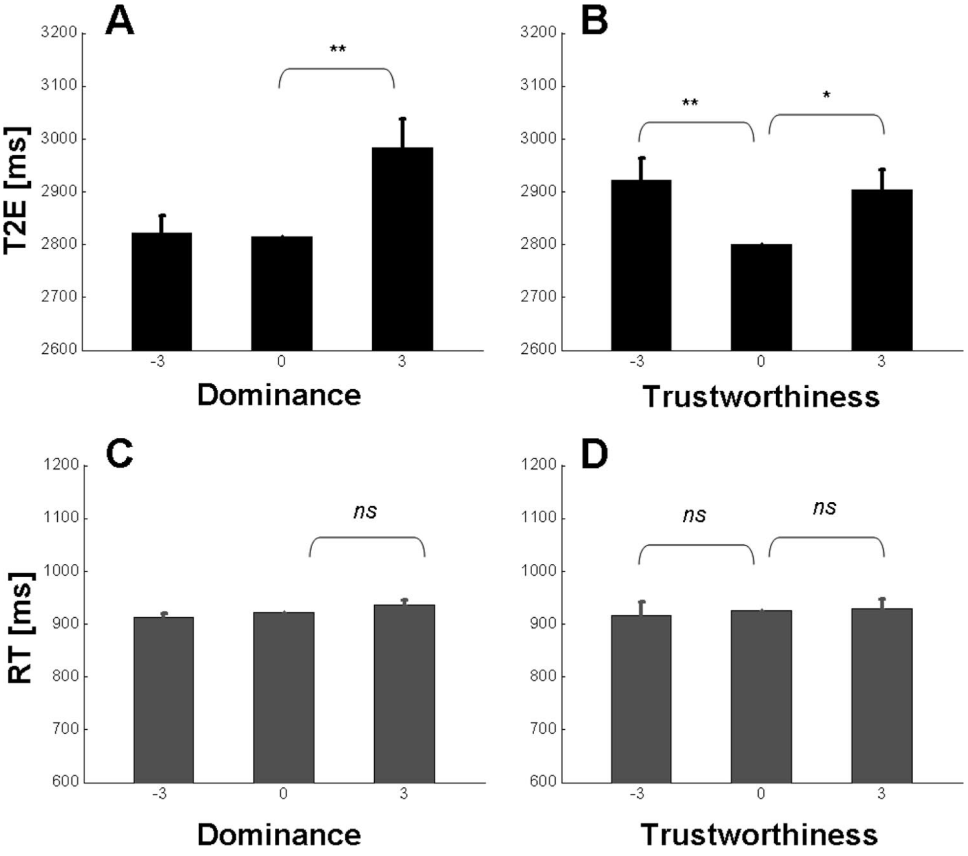

Figure 3. The results of Experiment 2. Each plot shows the average time to emerge (T2E) of subjects ($N = 23$) plotted against dominance (A, C) and trustworthiness (B, D) of the suppressed faces. The impact of each social dimension on responses is plotted separately (collapsing on the other dimension). Top (A, B) and bottom (C, D) panels correspond to suppressed and visible faces, respectively. Error bars = standard error of mean difference between the specific condition and neutral face. *ns*: $p > .05$. $^{*} p < .05$. $^{**} p < .01$.

similar trend for trust and visibility speak against a general threat-induced slowing of motor responses.

The reaction times in the visible condition (Figures 3C and 3D) were much faster (~2000 ms) than in the suppressed condition (Figures 3A and B), raising the possibility of a floor effect in the visible condition. However, we note that the purpose of Experiment 2 was to rule out processes that affect response times *after* breaking through suppression. The assumption of this experiment was that the processes after breakthrough (in the suppressed condition) are identical to the processes triggered by face presentation in the visible condition. As such, if social modulation does not affect reaction times to visible target (even if by a floor effect), it is not easy to see why it should affect post-breakthrough processes in the suppressed condition differently. To dismiss the results as a floor effect would be equivalent to abandoning the starting assumption of the experiment. We believe this is not plausible because the task was a simple left/right spatial localization in both conditions and did not require any detailed assessment of the faces in either condition.

The results of Experiment 2 show that the variability in T2E due to social evaluation was not explained away by post-breakthrough processes. They do not, however, serve to dismiss the possibility of any conscious behavioral effects arising from social evaluation of faces in general. We note here that we are not arguing for exclusivity of social evaluation effects for unconscious perception but rather to their existence in our experimental paradigm.

Experiment 3

Relating T2E From Suppression to Individual Differences in Personality Traits

People are differently inclined to trust others. Some can easily leave their lives in other people's hands, and some would shudder at imagining it. Similarly, some of people find the slightest social challenges unbearable and immediately submit while others thrive and blossom in the face of domineering opposition. These trait variations between individuals may indeed have correlates in face perception. Consistent with this proposition, when an affective and a neutral face compete for initial dominance in binocular rivalry, high trait-anxious observers more frequently report initial perception of negative expressions (i.e., fearful and angry faces; Gray, Adams, & Garner, 2009). This suggests that the observers' personality traits and preconscious face processing (i.e., neural events *before* the initial dominance) may be closely linked. To test the hypothesis that preconscious perception of trust and dominance are linked to the perceiver's personality traits, participants from the preceding experiments were asked to complete three validated questionnaires to assess propensity to trust, frequency of submissive behaviors, and social anxiety. The questionnaires were specifically selected to assess individual differences in personality traits that may relate to the social evaluation of faces along the axes of trust and dominance. We then tested whether there were any correlations between these self-reports and individual differences in behavior in our psychophysical experiments.

Description of the Questionnaires

Submissive Behavior Scale (SBS). A self-report 16-item measure of general submissive social behavior was adapted from a scale developed by Buss and Craik (1986). Participants are required to indicate how often they behave in a variety of submissive ways; scores range from 0 to 64, with higher scores indicating more frequent submissive behavior. The SBS shows internal validity and test–retest reliability (Allan & Gilbert, 1997). Cronbach's α in our sample was .92.

Propensity to Trust Survey (PTS; trust items only). Seven items loading most heavily onto the trust factor were taken from the full PTS, a 21-item questionnaire used to assess two dimensions: interpersonal trust and trustworthiness. Participants rate items according to the extent to which they consider the items descriptive of them; higher scores indicate a greater inclination to trust others. The full PTS has been shown to have good internal reliability and external validity (Evans & Revelle, 2008). Cronbach's α in our sample was .71.

Fear of Negative Evaluation (FNEB; brief version). The FNEB is a commonly used 12-item assessment of social anxiety. Participants are required to indicate how characteristic of them certain social anxiety behaviors are. Scores range from 12 to 60, with higher scores indicating elevated levels of social anxiety. The FNEB has been shown to be reliable (Leary, 1983) and has been validated both with patients (Collins, Westra, Dozois, & Stewart, 2005) and nonclinical samples (Watson & Friend, 1969). Cronbach's α in our sample was .76.

Method

Participants. Approximately 4 months after completion of the behavioral testing sessions, participants from previous experiments were invited via e-mail to complete the three questionnaires online; 22 participants responded. In order to increase our sample size for the correlational analysis, we conducted an abridged version of Experiment 1 with 28 additional participants, reaching a total of 50 participants.

Stimuli and procedure. This experiment was identical to the suppressed condition of Experiment 2. The nine face stimuli used in Experiment 2 were employed. Participants located the face relative (left vs right) to the fixation point. Each participant completed a total of 288 trials (eight blocks of 36 trials each), with each of the nine face versions used in this experiment presented a total of 32 times (four repetitions in each block). The participants in this experiment completed the questionnaires via e-mail about 1 week after taking the psychophysical experiment.

For each participant who completed the questionnaires, we calculated two psychophysical measures of preconscious social evaluation using the T2E data:

$$\text{Dominance avoidance} = \text{T2E}_{+3\text{dom}} - \text{T2E}_{\text{neutral}},$$

$$\text{Untrust avoidance} = \text{T2E}_{-3\text{trust}} - \text{T2E}_{\text{neutral}}$$

These measures were constructed from suppressed trials independent of the experiment in which the participant had taken part. We then tested whether the scores from the self-report questionnaires were correlated with these psychophysical measures of preconscious perception of dominance and trust.

Results

We first examined if the data from the newly acquired group of participants confirmed our earlier results. A two-way repeated-measure ANOVA (three levels of dominance and three levels of trust) revealed a significant main effect of dominance, $F(2, 54) = 8.85$, $p < .001$, partial eta squared = .247. The main effect of trust was marginally away from significance, $F(3, 54) = 2.76$, $p = .072$, partial eta squared = .09.The interaction between the two factors was not significant, $F(4, 108) = 2.50$, $p > .05$, partial eta squared = .08. Similar to Experiment 1, the quadratic component of variance was marginally significant for trust, $F(1, 27) = 3.88$, $p = .059$; partial eta squared = .126. Post hoc comparisons showed that T2E was significantly longer for most dominant versus neutral faces, $t(27) = 6.83$, $p < .001$, Cohen's $d = 2.6$, effect size correlation $r = .79$, and most untrustworthy versus neutral faces, $t(27) = 6.07$, $p < .001$, Cohen's $d = 2.33$, effect size correlation $r = .75$. Having replicated and confirmed the results of Experiments 1 and 2 in a new group of participants, we moved on to examine the role of individual differences in personality traits and behavioral performance in our experiment.

Self-reported submissive behavior was positively correlated with dominance avoidance, Pearson's $r = .29$; $p = .04$; confidence interval (CI) [0.013, 0.52] based on Fisher r-to-z transformation (Figure 4A). The more dominant faces remained suppressed for longer time in more submissive participants. Submissiveness did not predict the variation in avoidance of untrustworthy faces, $p > .2$ (Figure 4B). There was a strong inverse correlation between self-reported propensity to trust and both dominance avoidance, Pearson's $r = -.39$; $p = .005$, CI [-0.602, -0.126] (Figure 4C), and untrust avoidance, Pearson's $r = -.387$; $p = .005$, CI [-0.60, -0.122] (Figure 4D). Suspicious people (i.e., those reporting that they were unlikely to trust others easily) were slower to become aware of the masked untrustworthy as well as the dominant faces relative to the neutral faces. The degree of social anxiety, as measured by fear of negative evaluation, did not relate significantly to either of the behavioral effects ($p > .16$, Figure 4E; $p > .69$, Figure 4F). Inspection of the data revealed that the dominance avoidance index was distorted by outliers, and so four outliers (defined as more than 2 standard deviations from the mean) were replaced with the next nearest value plus 1 (procedure recommended for treatment of outliers by (Field, 2005). Outlier replacement did not significantly alter the magnitude, direction, or significance level of correlations.

So far, behavioral data for preconscious perception of trust showed a quadratic trend: the shortest T2E was observed for neutral faces while both trustworthy and untrustworthy faces took longer to emerge into awareness than neutral faces did. Trust avoidance ($\text{T2E}_{+3\text{trust}} - \text{T2E}_{\text{neutral}}$) did not show any significant correlation with any of the

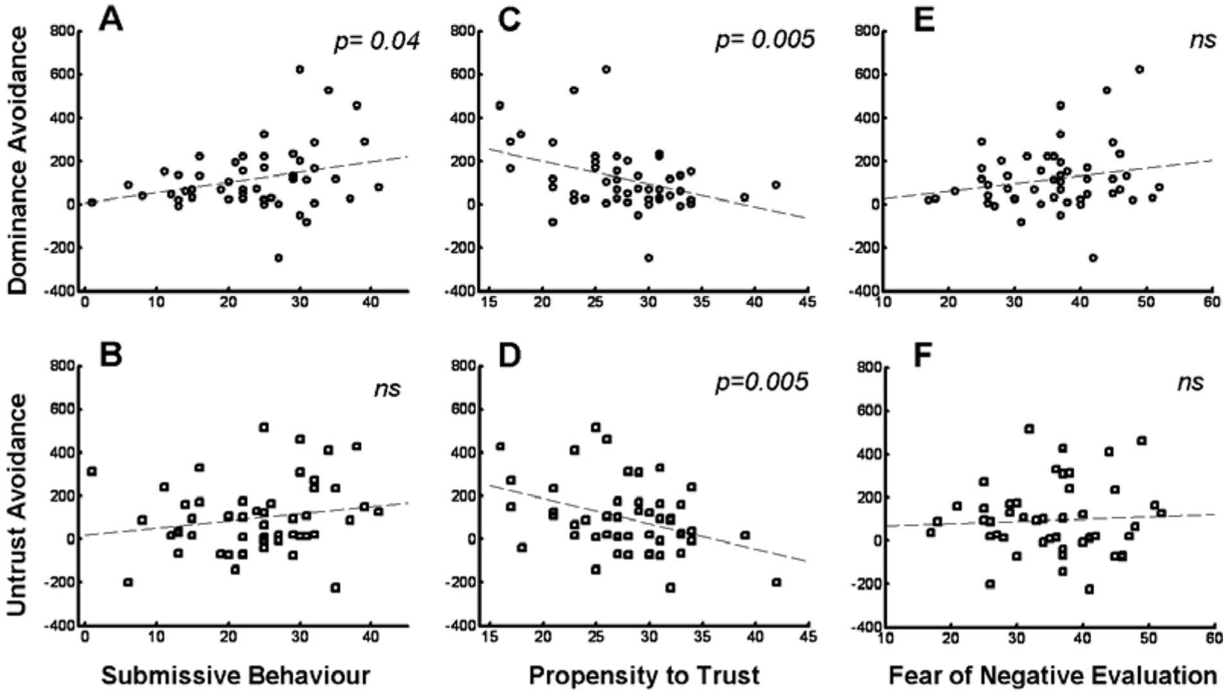

Figure 4. The relationship between self-reported personality traits of participants (x axis) and psychophysical performance (y axis) (*N* = *50*). Each plot shows the size of dominance avoidance effect (A, C, E) or the size of untrust avoidance effect (B, D, F) plotted against scores of personality trait questionnaires. Each subject is represented by a data point, and dashed lines represent lines of best fit.

three personality traits (all *p*s $> .2$; Pearson's $r < .19$). Similar negative results were found for nondominant ($\mathrm{T2E}_{-3\mathrm{dom}} - \mathrm{T2E}_{\mathrm{neutral}}$) condition (all *p*s $> .2$; Pearson's $r < 0.1$). We emphasize, however, that our main hypothesis connecting personality traits to preconscious social evaluation was mainly concerned with threatening (untrustworthy, dominant) faces that correspond to the main results described in this section (Figure 4).

These results suggest that self-reported social trait characteristics are related to the individuals' preconscious perception of socially relevant visual information. However, some cautionary notes must be made before moving on to the discussion. Within the psychophysical task, our two measures of preconscious avoidance (i.e., dominance and untrustworthiness) were clearly correlated (Pearson's $r = .42$; $p = .002$). Moreover, a strong, hitherto-unreported negative correlation was observed between SBS and PTS scores (Pearson's $r = -.59$; $p < .001$): participants who reported being more open to trusting others also reported less likelihood of submissive behavior. The implications of this finding that, to our knowledge, has not been reported before are beyond the scope of the current work. However, these two internal correlations (one between our two psychophysical measures and the other between our main personality measures) complicate the interpretation of how the observer-specific traits relate to preconscious perception, which is, indeed, a very interesting question for future research. Nonetheless, the main message from our findings—that observer-specific-traits *do* relate to preconscious social evaluations—is not affected by the internal relationships within each method for at least two reasons: (a) the psychophysical task was an incidental (left/right) judgment that was entirely orthogonal to the social labels assigned to the faces by the model (Oosterhof & Todorov, 2008) that generated the faces, and (b) the questionnaires responses were obtained with substantial (at least 1 week) time difference from the psychophysical testing.

Experiment 4

Conscious Ratings of Trust and Dominance

Throughout the experiments reported, we used computer-generated face stimuli for which trust and dominance were defined using a recent quantitative model that maps facial features onto evaluation of these two social labels (Oosterhof & Todorov, 2008). The validity of this mapping (i.e., whether, for example, a dominant face generated by the model is actually perceived and rated as dominant by human observers) has previously been tested by Oosterhof and Todorov in 2008 and in a subsequent study (Todorov, 2011). Nonetheless, we also conducted an ancillary experiment to provide independent evidence for validity of the social labels assigned by the model for the stimuli employed in our experiment. We asked a new (much larger) group of participants to view the computer-generated faces consciously and rate them on the two dimensions of trust and dominance.

Method

Participants. A total of 138 participants (87 women; age range = 17–73 years; mean age = 23.9 years) took part in this experiment. All participants had normal or corrected-to-normal vision. Participants gave written informed consent, and the experiment was approved by the local research ethics committee.

Stimuli and procedure. Programming software was identical to that used in Experiment 1. Stimuli were presented on a 19-in. Dell 60-Hz LCD monitor (Dell, Inc., Round Rock, TX). Because faces were presented consciously in this experiment, we did not use the stereoscope. In each trial, one face (6.3° height by 3.9° width in visual angle) was presented centrally for either 50 or 500 ms. After presentation, the participant used the mouse buttons to rate the face on one of the two dimensions (trust or dominance, decided by block) along a 7-point Likert scale. The face stimulus in each trial was pseudorandomly chosen from the 3 (trust) × 3 (dominance) space (see Method section in Experiment 2). Each of the nine combinations of trust and dominance was tested six times, giving rise to a total of 54 trials for the trust-rating block and 54 trials for the dominance-rating block. For each trial, face identity was randomly sampled from 10 possible options. For data analysis, the results were collapsed across identities.

Results and Discussion

Consistent with previous reports (Oosterhof & Todorov, 2008), trustworthiness generated by the model reliably predicted the trustworthiness judgments reported by the ratings (see Table 1); two-way repeated-measures ANOVA with three levels for trust and two levels for exposure duration showed a main effect of trustworthiness, $F(2, 274) = 271.41$, $p < .001$; partial eta squared = .665. There was no main effect of exposure duration, $F(1, 137) = 1.06$, $p = .31$, partial eta squared = .008. A significant interaction between exposure duration and trust level, $F(2, 274) = 6.51$, $p = .002$, partial eta squared = 0.045, was found such that untrustworthy faces were rated as more untrustworthy after 500-ms exposure than after 50-ms exposure, $t(137) = 3.21$, $p = .006$, corrected for multiple comparisons, Cohen's $d = 0.54$, effect size correlation $r = .26$. Trust ratings did not differ between 50-ms and 500-ms exposure times for either neutral, $t(137) = 0.10$, $p = .92$, or trustworthy, $t(137) = 0.95$, $p = .35$, faces.

Similarly, model-based dominance scores of faces predicted the dominance rating attributed by the participants (see Table 2); two-way repeated measures ANOVA with three levels for dominance and two levels for exposure duration showed a main effect of dominance, $F(2, 274) = 379.79$, $p < .001$, partial eta squared = .73. Dominance ratings also were significantly higher after 500-ms exposures than after 50-ms exposures: main effect of duration, $F(1, 137) = 6.17$, $p = .01$, partial eta squared = .043. An interaction between exposure duration and dominance level was also seen, $F(2, 274) = 14.58$, $p < .001$, partial eta squared = .098, such that dominance ratings were higher after 500 ms than 50 ms for both neutral and dominant faces, $t(137) = 5.10$, $p < .01$, corrected for multiple comparisons, Cohen's $d = 0.87$; effect size correlation $r = .39$, but did not differ significantly for neutral, $t(137) = 2.02$, $p = .14$, or submissive faces, $t(137) = 1.99$, $p = .14$.

Table 1
Results of Experiment 4: Means and Standard Deviations of Trust Ratings for Each Exposure Duration

	Trust level predicted by the model					
	Untrustworthy (−3 *SD*s)		Neutral		Trustworthy (+3 *SD*s)	
Exposure duration	*M*	*SD*	*M*	*SD*	*M*	*SD*
50 ms	3.83	1.21	4.99	1.10	5.87	1.21
500 ms	3.57	1.28	4.99	1.17	5.94	1.28

Table 2
Results of Experiment 4: Means and Standard Deviations of Dominance Ratings for Each Exposure Duration

	Dominance level predicted by the model					
	Submissive (−3 *SD*s)		Neutral		Dominant (+3 *SD*s)	
Exposure duration	*M*	*SD*	*M*	*SD*	*M*	*SD*
50 ms	3.82	1.03	4.62	0.95	5.70	1.02
500 ms	3.67	1.08	4.76	0.92	6.06	1.06

These results confirm that the model we used indeed generated face stimuli that accurately mapped the participants' conscious perception of trust and dominance. As such, the results validate in the United Kingdom the model for social evaluation of faces (Oosterhof & Todorov, 2008) that had previously been only tested in the United States. With these results in hand, we are confident that the observers' social evaluation of the faces is indeed aligned with the predictions of the model that generated the faces.

These results, however, cannot tell us whether invisible faces used in Experiments 1–3 could be evaluated along these social dimensions or not. As such, whether participants would be able to determine the trustworthiness or dominance of an invisible face remains an open question. The answer to this question would require an experiment in which the explicit focus of inquiry is the social evaluation of the invisible face. Our method (Experiments 1–3), however, relied on the T2E, which is an implicit measure of preconscious perception. This latter consideration relates to another critical question: are participants' personality traits more closely tied to the implicit preconscious (via differential time to emerge through CFS) or the explicit conscious evaluation (via direct rating) of the faces? Future research in social evaluation of faces could address this interesting question.

The interactions between exposure duration and trust/dominance judgments showed that allowing more time for inspection exaggerated the social evaluation of the most threatening faces. These data are consistent with previous reports that social evaluation is efficient enough to assess faces that have been viewed only for 50 ms (Todorov, Said, et al., 2008), but they also suggest that evaluation of the particularly threatening faces may be modulated by another, slower process of postperceptual appraisal (also see Tsuchiya, Moradi, Felsen, Yamazaki, & Adolphs, 2009). This issue is outside the focus of this article but also presents a potentially useful avenue for future research.

General Discussion

The results of three experiments showed that social evaluation of faces along the trust and dominance axes (Oosterhof & Todorov, 2008) extends to preconscious stages of perception. In all three experiments, dominant faces that were masked by intraocular suppression took significantly longer to break through suppression than neutral faces. Similarly, untrustworthy and trustworthy faces took longer to break through suppression than neutral faces. These results were consistent in all three experiments.

Experiment 2 showed that speeded responses to consciously perceived faces did not show a similar modulation by social labels, ruling out the possibility that a generalized slowing of motor responses to faces *after* conscious perception may be responsible for the main findings. Finally, we showed that the preconscious perception of trust and dominance could be meaningfully related to the perceiver's relevant personality traits. Participants who rated themselves more likely to trust others showed less preconscious aversion from the untrustworthy (or the dominant) compared with the neutral faces. Those who reported stronger tendencies for submissive social behavior showed a stronger preconscious aversion to the dominant than to the neutral faces. The meaningful direction of the correlations between valence-dependent prolongations of suppression and the participants' self-reported personality traits also ruled out the possibility of a generalized slowing of preconscious processing due to some unknown common physical difference between neutral and dominant/untrustworthy faces. Personality traits were not predictive of the variability in the prolonged suppression of the trustworthy (vs neutral) faces (see later discussion). These results clearly demonstrated the existence of preconscious social evaluation of faces that depends on the observer's personality characteristics. Such preconscious social evaluation of faces is consistent with some very recent reports that showed that directly gazing faces break through interocular suppression faster than faces with averted gaze (Stein, Senju, Peelen, & Sterzer, 2011).

Overall, the direction of our results was opposite to the predictions of the evolutionary vigilance hypothesis (Morris et al., 1999, 2001; Vuilleumier, 2005). That hypothesis argues that preconscious face processing is mainly concerned with increasing the chances of survival by contributing to rapid coarse detection of threats; that is, more dominant and less trustworthy faces should have reached awareness first, which was the opposite of what we found. Moreover, the vigilance hypothesis predicts that submissive or untrusting individuals should be more sensitive to dominant or untrustworthy faces presumably because these types of faces should be motivationally more salient or relevant to them. The individual difference analysis also rejected this prediction by revealing that the more submissive or untrusting participants took longer to become aware of the dominant or untrustworthy (vs neutral) faces.

A possible reason for this discrepancy may be found by considering the different implications of different categories of negative stimuli. A fearful face can be thought of as an indicator of nearby threat but is not itself a direct threat. Upon detecting a fearful face, perhaps a reasonable strategy would be to try to spot the actual source of the threat. Indeed, a number of studies have shown that fearful cues enhance visual contrast sensitivity (Phelps, Ling, & Carrasco, 2006) and visual search efficiency (Krusemark & Li, 2011). On the other hand, an angry face is much less ambiguous in directly showing the source of threat. Indeed, direction of gaze of fearful and angry faces modulates the response time (R. B. Adams & Kleck, 2003) and amygdala response (R. B. Adams, Gordon, Baird, Ambady, & Kleck, 2003) to these negative stimuli: faster reaction times and greater amygdala response are observed in response to "angry faces with direct gaze" (R. B. Adams et al., 2003, p. 1536) and "fearful faces with averted gaze" (R. B. Adams et al., 2003, p. 1536). Reaction times and brain response to directly gazing fearful faces—which break through CFS faster than neutral faces (Yang et al., 2007)—are significantly different from both direct-gazing angry and gaze-averted fearful faces (R. B. Adams & Kleck, 2003; R. B. Adams et al., 2003). We suggest that the dominant and untrustworthy faces used here are more likely to be direct sources of threat, similar to directly gazing angry faces.

Several studies have shown an advantage for detecting a negative (i.e., threatening, angry) oddball face among positive (e.g., happy) distractor faces. This face-in-the-crowd effect has been interpreted as a processing advantage for threat (Hansen & Hansen, 1988). A number of later studies challenged this interpretation and some alternative explanations (e.g., more efficient [happy] distractor exclusion) have been proposed for this effect (Frischen, Eastwood, & Smilek, 2008). Direct comparison of visual search (in these previous experiments) and breaking through CFS (as here) is difficult for a number of reasons. For example, there are no distractors in CFS paradigms, and the spatial attentional component of the face detection task is more relevant to visual search. However, the originally claimed superiority of anger in face-in-the-crowd effect predicts that directly gazing angry faces should break through CFS faster than happy faces. This prediction is also consistent with the vigilance theory. Our account, however, favors the opposite. Fortunately, a very recent study (Stein & Sterzer, in press) has tested this prediction using schematic angry and happy faces. The results show that angry faces take *longer* than happy ones to break through CFS. Nonetheless, it is important to note that while Stein and Sterzer's results confirm our prediction, they also raise further issues. Their results showed that the difference between angry and happy faces was explained by the difference between configural organization of mouth curvature and face contour between the two expressions. This led the authors to argue for a predominantly perceptual basis for the findings such as face-in-the-crowd effect, a conclusion also supported by other studies (Coelho, Cloete, & Wallis, 2010; Horstmann, Becker, Bergmann, & Burghaus, 2010) that examined the role of low level features in face-in-the-crowd effect directly.

Behavioral fear response to direct threat consists of two opposite extremes. At one extreme, direct threat may induce an active fear response characterized by fight/flight that is mediated by the amygdala, the cholinergic basal forebrain, and its neuromodulatory effect on cortical arousal. At the other extreme, encounter with direct threat may induce a passive response characterized by *freezing,* mediated by the amygdala, the brainstem, and the basal forebrain system (Pape, 2010). Freezing behavior has been extensively studied in rodents and nonhuman primates. But research on freezing induced by social threats has only recently started to be studied in humans (Roelofs, Hagenaars, & Stins, 2010; Terburg, Hooiveld, Aarts, Kenemans, & van Honk, 2011). One study showed that female observers' body sway and heart rate decreased when the women were viewing angry (vs neutral) faces, and this freezing-like response is correlated with state anxiety (Roelofs et al., 2010). In another study, observers were asked to divert their gaze from a briefly presented angry, happy, or neutral face. Use of forward and backward masking ensured that observers were not aware of the face's emotional valence. Saccadic latencies for diverting gaze away from the faces were longer for angry compared with neutral faces, and this effect correlated with the observers' trait dominance measured via a separate self-report questionnaire (Terburg et al., 2011).

Consistent with the previously described studies, and following from our earlier notion that directly gazing angry faces and dominant or untrusted faces may be treated similarly at an unconscious level, we suggest that the prolonged preconscious perception of dominant and untrustworthy faces reported here may in fact be the result of a passive fear response, leading to slowed visual perception. Our individual difference findings are also consistent with this notion: freezing (i.e., preconscious fear response) was stronger in the more submissive, untrusting personality types who were presumably more likely to take the threat more seriously.

A possible neuronal scenario underlying these findings is that the unconscious exposure to social threat (i.e., a dominant untrustworthy face) may trigger a passive fear response initiated by the amygdala, which, in turn, may mediate the reduction of neuromodulatory influence of the basal forebrain system on cerebral cortex, leading to reduced cortical arousal and prolonged suppression of threat-related stimuli (Pape, 2010). This view recognizes a role for the cortex in prolonged suppression of social threats, but it borrows its main premise from the subcortical fear detection theory (see introduction). However, a number of recent studies have challenged this notion by showing that amygdala responses to fear are not automatic but depend on attentional load (Pessoa, McKenna, Gutierrez, & Ungerleider, 2002; Pessoa, Padmala, & Morland, 2005). Moreover, rapid implicit detection of fear-related stimuli is intact in a patient with bilateral amygdala lesions (Tsuchiya et al., 2009). These studies raise the possibility that previous work may have underestimated the role of cortical structures in preconscious social and emotional perception. Taken together with our findings, we predict that the human amygdala and the visual cortex response to suppressed faces should distinguish between preconscious dominant/untrustworthy and neutral faces *in opposite directions*: preconscious dominant/untrustworthy faces should elicit stronger responses in amygdala and weaker responses in the visual cortex. These opposite effects should correlate with observer personality traits. However, based on the contrast between findings about fear and anger described earlier, we predict that patients with bilateral amygdala lesions (Adolphs, Tranel, Damasio, & Damasio, 1994) will not show the prolonged suppression for dominant/untrustworthy faces. This latter prediction is the exact opposite of the findings from suppression of fear by CFS in patients with amygdala lesions (Tsuchiya et al., 2009). Future research on the neural basis of preconscious emotional and social perception will allow investigators to distinguish between these alternatives.

In all three experiments, we observed a U-shaped pattern for the preconscious perception of trust: both trustworthy and untrustworthy faces took longer to emerge into participants' awareness compared with neutral faces. These behavioral findings parallel a number of recent works that have reported a similar quadratic pattern of human brain response—most predominantly the human amygdala—to trustworthiness of faces (Said, Baron, & Todorov, 2009; Todorov, Baron, & Oosterhof, 2008; Todorov, Said, Oosterhof, & Engell, 2011). It has recently been argued (Said, Dotsch, & Todorov, 2010) that neuronal responses in the human amygdala and fusiform face area may be driven by the distance from the average face. Face typicality covaries with trust but not with dominance (Todorov et al., 2011). Therefore, an alternative interpretation of the U-shape pattern of findings reported here could be that the typical (average face) emerges faster than the atypical trustworthy and untrustworthy faces. However, this interpretation alone cannot account for why individual differences are predictive of variation in T2E for untrustworthy but not for trustworthy faces. Together, these opposite considerations suggest that preconscious processing of trustworthy and untrustworthy faces may be driven by different underlying mechanisms. For future research, it would be interesting to compare CFS responses to faces that are matched on distance with average faces that are mismatched on social valence using stimuli similar to Said et al. (2010).

Finally, our results disclose and emphasize the importance of individual differences in personality traits and their relevance to face perception. Highlighting the relevance of variability in observer-specific traits for face perception may provide clues for explaining some of the discrepancies in unconscious high-level face perception literature. The seemingly contradictory results from previous work on high-level face adaptation (Adams et al., 2010; Amihai et al., in press; Moradi et al., 2005) may be explained by individual differences between their participants arising from personality traits as well as structural brain differences (Kanai & Rees, 2011). The relatively small sample sizes employed in psychophysical studies of adaptation may have exacerbated this situation. This issue is of special relevance because traditionally, the holy grail of psychophysical research has been to identify the principles governing perception that are independent of observers. This goal is often pursued by focusing on extensive data collection from small numbers of participants, a practice based on the assumption that between-observer differences must be of little relevance. Our results, along with some other recent findings (Gray et al., 2009), caution against such simplifying assumptions.

References

Adams, R. B., Jr., Gordon, H. L., Baird, A. A., Ambady, N., & Kleck, R. E. (2003, June 23). Effects of gaze on amygdala sensitivity to anger and fear faces. *Science, 300,* 1536. doi:10.1126/science.1082244

Adams, R. B., Jr., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science, 14,* 644–647. doi:10.1046/j.0956-7976.2003.psci_1479.x

Adams, W. J., Gray, K. L., Garner, M., & Graf, E. W. (2010). High-level face adaptation without awareness. *Psychological Science, 21,* 205–210. doi:10.1177/0956797609359508

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994, December 15). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372,* 669–672. doi: 10.1038/372669a0

Allan, S., & Gilbert, P. (1997). Submissive behaviour and psychopathology. *British Journal of Clinical Psychology, 36,* 467–488. doi:10.1111/j.2044-8260.1997.tb01255.x

Alpers, G. W., & Gerdes, A. B. (2007). Here is looking at you: Emotional faces predominate in binocular rivalry. *Emotion, 7,* 495–506. doi: 10.1037/1528-3542.7.3.495

Amihai, I., Deouell, L., & Bentin, S. (in press). Conscious awareness is necessary for processing race and gender information from faces. *Consciousness and Cognition.*

Bahrami, B., Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2008a). Spatial attention can modulate unconscious orientation processing. *Perception, 37,* 1520–1528. doi:10.1068/p5999

Bahrami, B., Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2008b). Unconscious orientation processing depends on perceptual load. *Journal of Vision, 8,* 12. doi:10.1167/8.3.12

Bahrami, B., Lavie, N., & Rees, G. (2007). Attentional load modulates responses of human primary visual cortex to invisible stimuli. *Current Biology, 17,* 509–513. doi:10.1016/j.cub.2007.01.070

Bahrami, B., Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., & Rees, G. (2010). Unconscious numerical priming despite interocular suppression. *Psychological Science, 21,* 224–233. doi:10.1177/0956797609360664

Buss, D. M., & Craik, K. H. (1986). Acts, dispositions, and clinical-assessment: The psychopathology of everyday conduct. *Clinical Psychology Review, 6,* 387–406. doi:10.1016/0272-7358(86)90028-0

Coelho, C. M., Cloete, S., & Wallis, G. (2010). The face-in-the-crowd effect: When angry faces are just cross(es). *Journal of Vision, 10,* 1–14.

Collins, K. A., Westra, H. A., Dozois, D. J., & Stewart, S. H. (2005). The validity of the brief version of the Fear of Negative Evaluation Scale. *Journal of Anxiety Disorders, 19,* 345–359. doi:10.1016/j.janxdis.2004.02.003

Evans, A. M., & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality, 42,* 1585–1593. doi:10.1016/j.jrp.2008.07.011

Field, A. (2005). *Discovering statistics using SPSS: (and sex, drugs and rock ‘n’ roll)* (2nd ed.) London, England: Sage.

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological Bulletin, 134,* 662–676. doi:10.1037/0033-2909.134.5.662

Gray, K. L., Adams, W. J., & Garner, M. (2009). The influence of anxiety on the initial selection of emotional faces presented in binocular rivalry. *Cognition, 113,* 105–110. doi:10.1016/j.cognition.2009.06.009

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54,* 917–924. doi:10.1037/0022-3514.54.6.917

Horstmann, G., Becker, S. I., Bergmann, S., & Burghaus, L. (2010). A reversal of the search asymmetry favouring negative schematic faces. *Visual Cognition, 18,* 981–1016. doi:10.1080/13506280903435709

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *PNAS: Proceedings of the National Academy of Science of the United States of America, 103,* 17048–17052. doi:10.1073/pnas.0605678103

Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological Science, 18,* 349–355. doi:10.1111/j.1467-9280.2007.01902.x

Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: Dissociating subsystems for facial-information processing. *Current Biology, 16,* 2023–2029. doi:10.1016/j.cub.2006.08.084

Jiang, Y., Zhou, K., & He, S. (2007). Human visual cortex responds to invisible chromatic flicker. *Nature Neuroscience, 10,* 657–662. doi:10.1038/nn1879

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat Rev Neurosci, 12,* 231–242. doi:10.1038/nrn3000

Krusemark, E. A., & Li, W. (2011). Do all threats work the same way? Divergent effects of fear and disgust on sensory perception and attention. *The Journal of Neuroscience, 31,* 3429–3434. doi:10.1523/JNEUROSCI.4394-10.2011

Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin, 9,* 371–375. doi:10.1177/0146167283093007

Moradi, F., Koch, C., & Shimojo, S. (2005). Face adaptation depends on seeing the face. *Neuron, 45,* 169–175. doi:10.1016/j.neuron.2004.12.018

Morris, J. S., DeGelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain, 124,* 1241–1252. doi:10.1093/brain/124.6.1241

Morris, J. S., Ohman, A., & Dolan, R. J. (1998, June 4). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393,* 467–470. doi:10.1038/30976

Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating “unseen” fear. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 96,* 1680–1685. doi:10.1073/pnas.96.4.1680

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 105,* 11087–11092. doi:10.1073/pnas.0805664105

Pape, H. C. (2010). Petrified or aroused with fear: The central amygdala takes the lead. *Neuron, 67,* 527–529. doi:10.1016/j.neuron.2010.08.009

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 99,* 11458–11463. doi:10.1073/pnas.172403899

Pessoa, L., Padmala, S., & Morland, T. (2005). Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *NeuroImage, 28,* 249–255. doi:10.1016/j.neuroimage.2005.05.048

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science, 17,* 292–299. doi:10.1111/j.1467-9280.2006.01701.x

Roelofs, K., Hagenaars, M. A., & Stins, J. (2010). Facing freeze: Social threat induces bodily freeze in humans. *Psychological Science, 21,* 1575–1581. doi:10.1177/0956797610384746

Said, C. P., Baron, S. G., & Todorov, A. (2009). Nonlinear amygdala response to face trustworthiness: Contributions of high and low spatial frequency information. *Journal of Cognitive Neuroscience, 21,* 519–528. doi:10.1162/jocn.2009.21041

Said, C. P., Dotsch, R., & Todorov, A. (2010). The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia, 48,* 3596–3605. doi:10.1016/j.neuropsychologia.2010.08.009

Shin, K., Stolte, M., & Chong, S. C. (2009). The effect of spatial attention on invisible stimuli. *Attention, Perception, & Psychophysics, 71,* 1507–1513. doi:10.3758/APP.71.7.1507

Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition, 119,* 307–311. doi:10.1016/j.cognition.2011.01.008

Stein, T., & Sterzer, P. (in press). Not just another face in the crowd: Detecting emotional schematic faces during continuous flash suppression. *Emotion.*

Sterzer, P., Jalkanen, L., & Rees, G. (2009). Electromagnetic responses to invisible face stimuli during binocular suppression. *NeuroImage., 46,* 803–808. doi:10.1016/j.neuroimage.2009.02.046

Sterzer, P., & Rees, G. (2008). A neural basis for percept stabilization in binocular rivalry. *Journal of Cognitive Neuroscience, 20,* 389–399. doi:10.1162/jocn.2008.20039

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychological Science, 21,* 349–354. doi:10.1177/0956797610362647

Terburg, D., Hooiveld, N., Aarts, H., Kenemans, J. L., & van Honk, J. (2011). Eye tracking unconscious face-to-face confrontations: Dominance motives prolong gaze to masked angry faces. *Psychological Science, 22,* 314–319. doi:10.1177/0956797611398492

Todorov, A. (2011). Evaluating faces on social dimensions. In A. Todorov, S. T. Fiske, & D. Prentice (Eds.), *Social neuroscience: Toward understanding the underpinnings of the social mind* (pp. 54–76). Oxford, England: Oxford University Press.

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience, 3,* 119–127. doi:10.1093/scan/nsn009

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005, June 10). Inferences of competence from faces predict election outcomes. *Science, 308,* 1623–1626. doi:10.1126/science.1110589

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Science, 12,* 455–460. doi:10.1016/j.tics.2008.10.001

Todorov, A., Said, C. P., Oosterhof, N. N., & Engell, A. D. (2011). Task-invariant brain responses to the social value of faces. *Journal of Cognitive Neuroscience.*

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8,* 1096–1101. doi:10.1038/nn1500

Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., & Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nature Neuroscience, 12,* 1224–1225. doi:10.1038/nn.2380

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Science, 9,* 585–594. doi:10.1016/j.tics.2005.10.011

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33,* 448-&. doi:10.1037/h0027806

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S. . . . Johnstone, T. (2004, December 15). Human amygdala responsivity to masked fearful eye whites. *Science, 306,* 2061. doi:10.1126/science.1103617

Yang, E., Hong, S. W., & Blake, R. (2010). Adaptation aftereffects to facial expressions suppressed from visual awareness. *Journal of Vison, 10,* 24. doi:10.1167/10.12.24

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion., 7,* 882–886. doi:10.1037/1528-3542.7.4.882

Yoon, K. L., Hong, S. W., Joormann, J., & Kang, P. (2009). Perception of facial expressions of emotion during binocular rivalry. *Emotion, 9,* 172–182. doi:10.1037/a0014714

Received April 3, 2011
Revision received January 23, 2012
Accepted January 30, 2012 ■