DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Separable Neural Components in the Processing of Black and White Faces

(Article begins on next page)

Running Head: SEPARABLE COMPONENTS OF SOCIAL PROCESSING

Separable Neural Components in the

Processing of Black and White Faces

William A. Cunningham[1], Marcia K. Johnson[1], Carol L. Raye[1]

J. Chris Gatenby[2], John C. Gore[2], & Mahzarin R. Banaji[3]

[1]Department of Psychology: Yale University

[2]Institute of Imaging Science: Vanderbilt University

[3]Department of Psychology: Harvard University

Corresponding author: William A. Cunningham, Department of Psychology, Yale University,
P.O. Box 208205, New Haven, Connecticut 06520. Phone: 203.432.9266. Fax: 203.432.7172
e-mail: william.cunningham@yale.edu. Phone (203)432-9266

To examine the neural components of automatic and controlled social evaluation, White participants viewed short (30ms) and long (525ms) duration Black and White faces during event-related functional magnetic resonance imaging (fMRI). At 30ms, greater activation in the amygdala – a brain region associated with emotion – was observed to Black than White faces. At 525ms, this difference was significantly reduced, and accompanied by Black>White activation in regions of frontal cortex associated with control/regulation. Furthermore, participants showing greater race bias on an indirect behavioral measure showed greater Black>White amygdala activation, and frontal activity predicted reduced Black>White differences in amygdala activity between the 30ms and 525ms conditions. These results provide evidence for neural distinctions between automatic and more controlled processing of social groups, and suggest that controlled processes may modulate automatic evaluation.

Separable Neural Components in the Processing of Black and White Faces

For almost 100 years, psychologists have studied attitudes and preferences by asking people to report on the good-bad attributes of people, things, and events (Eagly & Chaiken, 1993). However, recent evidence shows that people also spontaneously evaluate social objects along a good-bad dimension, without necessarily being aware that they are doing so (Bargh, Chaiken, Govender, & Pratto, 1992; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Given such findings, models of social attitudes suggest at least two modes of evaluation: one that involves conscious and controlled modes of thinking; and another that involves relatively automatic processes that operate without deliberate thought or sometimes even without conscious awareness (Greenwald & Banaji, 1995; Nisbitt & Wilson, 1977). Furthermore, an evaluation following more controlled evaluative processing may differ from an evaluation based only on relatively more automatic processing. In evaluations of social groups, for instance, many White participants show negativity toward Black, elderly, or foreigner compared with White, young, or American on indirect or implicit measures that tap automatic associations, yet often report unbiased attitudes on questionnaires that allow more controlled or conscious evaluations of the same groups (Cunningham, Nezlek, & Banaji, in press; Devine, 1989; Nosek, Banaji, & Greenwald, 2002).

Rather than absolute categories, automatic and controlled can be thought of as relative terms used as shorthand for referring to differences in the nature, number and/or complexity of the cognitive operations engaged (e.g., Johnson & Reeder, 1997). That is, automatic processes are said to require fewer cognitive resources, and involve less intent and/or conscious experience (see Bargh, 1996 for a review). Furthermore, cognitive operations that are thought to involve intentions, more complex operations, or the retrieval of more complex information take more

time than more automatic or less complex processes (e.g., Neely, 1977; Johnson, Kounios, &

Reeder, 1994). Therefore, in the current study, by manipulating the time that stimuli are

presented, we contrast conditions that vary the opportunity for conscious perception (e.g.,

Cheesman & Merikle, 1986; Marcel, 1983), and the opportunity for reflective or controlled

processing of the stimuli (Murphy, Monahan, & Zajonc, 1995; Murphy & Zajonc, 1993; von

Hippel & Hawkins, 1994). Differences in responses evoked by stimuli that are presented for a

very short interval presumably reflect differences arising from relatively automatic perceptual

and associative operations whereas differences in responses evoked by stimuli that are clearly

visible should be more likely to reflect controlled cognitive operations as well (see Draine &

Greenwald, 1998 regarding the use of briefly represented stimuli to investigate automatic

evaluative processes).

Here, we integrate behavioral work on automatic and controlled evaluation with an

investigation of the neural systems that may underlie these processes. Two primary questions are

posed: First, are common brain areas involved in both the automatic and controlled processing of

social group members, or do different brain areas contribute to these seemingly distinct

processes? Second, if different brain regions are recruited in relatively more controlled compared

with relatively more automatic processing of members of social groups, is there evidence that

more controlled processing can modulate the activity resulting from automatic processing?

Prior findings suggest that the amygdala is responsive to the emotionality of stimuli

(Davis & Whalen, 2001; Isenberg, Silbersweig, Engelien et al., 1999; LaBar, Gatenby, Gore,

LeDoux, & Phelps, 1998; LeDoux, 1996). Also, the amygdala responds to emotion (e.g.,

fearful>neutral facial expressions) whether stimuli are presented at durations long enough to be

consciously seen (Morris, Frith, Perrett et al., 1996) or more briefly (33ms and masked; Whalen

Rauch, Etcoff et al., 1998). Given the strong evidence for White participants' negative evaluations of Black compared to White individuals on measures of automatic evaluation (Cunningham et al., in press; Nosek et al., 2002; Rudman, Greenwald, Mellott, & Schwartz, 1999; Wittenbrink, Judd, & Park, 1997), one might expect greater amygdala activation for Black faces compared with White faces. To date, however, neuroimaging studies using presentations of Black and White faces have not found this expected pattern of greater amygdala activation to Black compared with White faces (Hart, Whalen, Shin et al., 2000; Phelps, O'Connor, Cunningham et al., 2000). Although Hart et al. showed amygdala activity to same race faces habituated more quickly than activity to other race faces, there was no overall peak difference in activation levels. In addition, even though Phelps et al. found amygdala activity to Black faces relative to White faces to be positively correlated with indirect measures of prejudice, there was no overall difference in amygdala activation to Black and White faces.

The failure to find this expected difference could be the result of conflict between automatic and more controlled processes. Presumably, automatic and controlled responses to fearful faces do not differ (both negative). In contrast, automatic and controlled evaluations of social groups do sometimes differ. If more positive controlled processing can moderate more negative automatic processing, this could account for the absence of clear differences in the amygdala response to Black and White faces in previous studies that have exclusively used stimuli which participants could see and reflect upon. That is, previous studies may have failed to detect differences in amygdala activity because of the simultaneous activation of discrepant positive (more reflective) and negative (more automatic) evaluation processes (Greenwald & Banaji, 1995; Johnson & Multhaup, 1992).

In this study, we directly compared two perceptual conditions that varied the opportunity for controlled processing of race by presenting photographs of Black and White faces either for very short durations (30ms; barely a flash on the screen) or at durations long enough to be clearly visible (525ms). Differences in activation to Black and White faces presented at short duration reflect relatively more automatic perceptual processing compared with differences in activation to faces that participants can clearly see, which allows for more opportunity to engage in reflective or controlled processing of faces.

Methods

Participants

Twenty White participants were paid for their participation. The data from four participants were omitted for excessive head movement (greater than 2mm in any direction) and from three other participants who when asked reported that they may have seen face-like stimuli in the 30ms presentations. The final 13 participants (4 females) had a mean age of 27.

Task

During fMRI, participants pressed one of two buttons with their right hand to indicate whether a visual stimulus appeared to the left or right of a fixation cross. Stimuli were forward projected using a projection panel and overhead projector onto a clear screen at the base of the MRI bore. The projection panel had a refresh rate of 15ms. From the participants' point of view, stimuli were abstract pictures, white squares, or emotionally neutral human faces. Six trial types were constructed to present short-duration Black faces, long-duration Black faces, short-duration White faces, long-duration White faces, and short and long-duration white squares used as filler trials. Short-duration stimuli were presented for 30ms, and long-duration stimuli for 525ms. All stimuli were preceded and followed by an abstract picture, which masked the 30ms images. The

duration of the second mask image held the overall amount of visual information constant – the

second abstract picture was presented for 525ms on short-duration trials and 30ms on long-

duration trials.[1,2] Thus, on short-duration trials, participants reported seeing and judging right/left

for an abstract picture (no participant included in the analysis reported seeing the 30ms faces). A

white cross appeared for 1300ms between trials. To ensure that faces were always separated by

12 seconds, five white square trials followed each face presentation, each randomly presented as

a short or long-duration trial such that the interval between faces for short and long duration

trials was, on average, identical. Four runs of data were collected. Each run contained 6

presentations of each critical trial type (short-Black, short-White, long-Black, and long-White)

presented in random order. The same Black and White faces (N = 8 each) were presented in both

short (30ms) and long-duration (525ms) conditions.

Scanning

A GE 1.5T MRI scanner at the Yale Magnetic Resonance Research Center was used for

scanning. To cover the frontal lobe and the majority of the temporal and parietal lobes, eighteen

coronal slices (slice thickness: 6mm, skip = 2 mm) were prescribed perpendicular to the AC-PC

line, with the ninth slice centered on the amygdala. Functional images were acquired using a

single shot gradient echoplanar pulse sequence (TE = 60ms, TR = 2000ms, in-plane resolution =

3.125 x 3.125 mm, matrix size = 64 x 64, and FOV = 20 x 20 cm).

Preprocessing

Data were corrected for slice acquisition time and motion using SPM99 (Friston, Holmes,

Worsley et al., 1995), then coregistered to in-plane anatomical images and transformed to

conform to the SPM99 standard T1 MNI brain interpolated to 3x6x3 mm. Functional data were

smoothed using a 9mm FWHM (full-width-half-maximum) kernel, and default SPM99 high and low pass filters were applied.

Group Contrast Analyses

For each subject, statistical contrast maps were generated using SPM99 from the average fMRI BOLD signal for Black compared with White faces for both the short and long-duration conditions. To characterize the neural response for short and long-duration presentations of Black and White faces, the fMRI signal reflecting brain activity to the Black face trials and to the White face trials was regressed onto a canonical hemodynamic response and a secondary time derivative that allowed for differences in hemodynamic onset. To generate group statistical contrast maps, a random effects analysis was conducted using individual subject contrast maps as input. For analyses of the amygdala, an anatomical mask was created to reduce the opportunity of Type I error in this a priori region of interest. For whole brain analyses, significant areas of activity were defined as those in which at least 13 contiguous voxels differed in activity to Black and White faces at a significance level greater than $p < .005$ ($t > 3.05$) in the short or long-duration condition. The choice of significance threshold and cluster size was determined by a monte carlo simulation that took into consideration the original acquisition voxel size, the resliced voxel size, and smoothing kernel to obtain a corrected alpha level of p $< .05$. Time courses were plotted from voxels that differed in either the 30ms or 525ms contrasts of Black and White faces using the SPM ROI Toolbox (Poldrack, nd).

Behavioral Measures

The Implicit Association Test (IAT; Greenwald, McGhee, & Schartz, 1998) and all self-report questions were administered on a computer after scanning. For the IAT, participants were presented a face or a word on each trial. In one block of trials, participants pressed one computer

key for Black faces and words with good meanings (e.g. fabulous, terrific), and another key for

White faces and words with bad meanings (e.g., horrible, revolting). In another block of trials,

participants classified the Black faces and bad words with one key and the White faces and good

words with another. Implicit attitude was defined as the average difference in response latency

between these two conditions, such that higher scores reflected more difficulty pairing Black

with good than Black with Bad. Faces used for the Implicit Association Test and the brain

imaging part of the study were identical. [3]

After the IAT, the Modern Racism Scale (a self-report scale measuring agreement with

questions of racial prejudice; McConahay, 1986), and the Motivation to Respond without

Prejudice Scale (a self-report scale measuring motivations to think and behave without prejudice;

Plant & Devine, 1998) were presented in random order on the computer. Participants responded

to the items on the self-report measures with a rating from 1 to 6 where 1 indicated that the

participant strongly disagreed with the presented statement and 6 indicated that the participant

strongly agreed with the presented statement. We used the internal motivation subscale which

measures participants' desires to think, feel, and respond without prejudice for personal reasons.

The Modern Racism and Internal Motivation to Respond Without Prejudice Scales were

combined into a single self-reported prejudice index. Higher scores reflected more positive

attitudes toward Black Americans and a personal desire to act without prejudice toward them.

For the prejudice index, we subtracted the midpoint of the scale such that positive scores indicate

agreement with non-prejudiced items and disagreement with prejudicial items, and negative

scores indicate the reverse.

In addition, an Implicit-Explicit Conflict index was created that conceptually reflects the

discrepancy between automatic and controlled attitudes. Attitude conflict was operationalized as

the product (see Botvinick, Braver, Barch, Carter, & Cohen, 2001) of the IAT and prejudice

index such that higher scores reflected greater attitude conflict. The more a participant both

reported positive attitudes on the self-report measures, and showed negativity toward Black

relative to White on the IAT, the higher their conflict score.

<div align="center">Results</div>

Behavioral Data

All participants (13 out of 13) disagreed with prejudiced statements, agreed with non-

prejudiced statements, and reported having personal motivation to respond without bias (see

Table 1). Yet, on average, participants showed automatic negative associations toward Black

relative to White faces on the IAT. Response times (851 ms) from blocks pairing Black faces and

bad words, and White faces and good words, were significantly shorter than response times

(1011 ms) from blocks pairing Black faces and good words and White faces and bad words, $\underline{t}(11)$

= 3.6, $\underline{p}$ < .01, indicating stronger associations between Black+bad/White+good than

Black+good/White+bad.

fMRI Data

We assessed the differential neural response to faces by contrasting the fMRI signal to

Black and White faces for the 30ms and 525ms conditions. The 30ms contrast showed a

significant difference in an area of right amygdala which extended into the ventral pallidum,

$t(12) = 4.26$, $p < .001$. As can be seen in Figure 1, relative to non-face trials (white squares and

abstract pictures), amygdala activity for 30ms presentations increased to Black faces, and

decreased to White faces. Thresholding more liberally indicated that this activation spread

additionally into more ventral parts of the amygdala. This pattern suggests that, at more

automatic levels of processing, more emotional processing occurs for Black than White faces.

Support for the suggestion that this differential amygdala activation is part of an automatic evaluation process comes from correlating the mean difference in fMRI signal in the amygdala voxels found to be significant ($p < .005$) in the 30ms Black>White contrast with the implicit attitude measure collected after scanning. To the extent that amygdala activity reflects processing associated with the relatively automatic evaluation of social groups, individual differences in this activity should correlate with individual difference measures assumed to measure automatic biases. This analysis indicated that amygdala activity was significantly correlated, $r(11) = .79$, $p < .01$, with participants' IAT scores. That is, the more implicit negativity toward Black relative to White that participants showed on the IAT, the greater their amygdala activity to Black relative to White faces in the 30ms condition during scanning.

Although participants showed a greater amygdala response to Black compared with White faces presented for 30ms, research has suggested that many White individuals are motivated to regulate or control unwanted feelings toward particular social groups. All participants in this study fell into this category – despite showing more positive associations to White than Black on the IAT, all participants reported internalized desires to respond without prejudice. Consistent with this regulatory motivation, and in contrast to the difference in amygdala response to Black versus White faces in the 30ms condition, we found no significant amygdala difference in the Black>White contrast for faces presented for 525ms. Moreover, a direct comparison of amygdala activation to Black compared with White faces in the short and long-duration conditions resulted in a significant interaction. As can be seen in Figure 1, compared to activations to the 30ms condition, the difference in activation to Black faces relative to White faces was significantly reduced in the 525ms condition when a presumably countervalanced, more positive, attitude could also be activated, $F(1,12) = 5.25$, $p < .05$.

At the same time, areas of increased activity to Black faces relative to White faces in the

525ms condition (Figure 2) were observed in right ventrolateral PFC (BA 47: $t(12) = 4.04$, $p <$

.005), right dorsolateral PFC (BA 9: $t(12) = 4.88$, $p < .001$), and anterior cingulate (BA 32: $t(12)$

$= 5.82$, $p < .001$). We compared the magnitude of the Black>White difference in amygdala

activation in the 30ms condition to the corresponding difference in the 525msec condition (i.e.,

$[B-W]_{30ms} - [B-W]_{525ms}$) to generate an index of amygdala modulation. Correlating this

modulation index with participant Black>White contrast maps (voxel by voxel) for the 525ms

condition indicated that the modulation of amygdala activation was associated with increases in

activation in dorsolateral prefrontal cortex ($t(12) = 3.07$, $p < .005$; MNI: 33, 48, 36) and anterior

cingulate ($t(12) = 2.98$, $p < .01$; 3, 6, 33). Consistent with previous research that has shown that

these areas are associated with regulation and executive function (Beauregard, Levesque, &

Bourgouin, 2001; Ochsner, Bunge, Gross, & Gabrieli, 2002; MacDonald, Cohen, Stenger, &

Carter, 2000), these results suggest that controlled processing can moderate, and even override,

activity that would otherwise arise from automatic processing (Moskowitz, Gollwitzer, Wasel, &

Schaal, 1999).

In previous work, we showed that evaluations marked by ambivalence are associated

with activity in ventrolateral prefrontal cortex (Cunningham, Johnson, Gatenby et al., 2003). To

examine the role of the prefrontal cortex in attitudinal ambivalence arising from differences

between automatic and controlled attitudes, we calculated a discrepancy index of attitudinal

conflict in the present study as the product between participants' automatic and reverse coded

self-report measures of race attitudes (Table 1). Correlating this index with the Black >White

activation to faces in the 525ms condition indicated that greater attitude ambivalence was

significantly correlated with greater activation in an area of ventrolateral prefrontal cortex

adjacent to that found in the 525ms Black>White comparison, $t(11) = 3.14$, $p < .005$; MNI: 42, 20, -9. Moreover, we predicted activation to faces in the 525ms condition from this discrepancy index and the IAT simultaneously. This analysis indicated that, after controlling for individual differences in self-reported versus automatic attitude discrepancy, the IAT was a significant predictor of amygdala activation, $t(11) = 3.87$, $p < .01$. That is, in the 30ms condition, the IAT correlated directly with Black>White amygdala activity, and in the 525ms condition the IAT correlated with Black>White amygdala activity controlling for discrepancies between indirect and self-report measures of racial attitudes.

Discussion

In summary, we found greater amygdala activation for Black than White faces when faces were presented for only 30ms (a duration at which all participants reported only seeing the mask stimulus that followed the face). This difference in amygdala activation was strongest for participants with higher degrees of racial bias on the Implicit Association Test. These results, combined with previous investigations of intergroup attitudes, suggest that implicit negative associations to a social group may result in an automatic emotional response when encountering members of that group. Yet, when participants had the opportunity to process Black and White faces for 525ms (and reported seeing the faces), we observed activity differences not in the amygdala, but in areas of PFC (BA 47 & 9) and anterior cingulate (BA 32) – areas associated with inhibition, conflict, and control (Beauregard et al., 2001; Ochsner et al., 2002; MacDonald et al., 2000). Furthermore, activation in the ventrolateral prefrontal cortex was correlated with attitudinal ambivalence and activations in dorsolateral prefrontal and anterior cingulate cortices were correlated with an index of the modulation of amygdala activity when participants had more opportunity to reflectively process the faces.

Such a pattern is consistent with a suggestion by Richeson, Baird, Gordon et al. (2003) that activation in dorsolateral prefrontal cortex and anterior cingulate is associated with attempts to control unwanted prejudicial responses to Black faces. Richeson et al. found that people who have the strongest race bias on the IAT (and thus the most to control) had the largest degree of activation in these regions. Importantly, the regions thought to underlie the control of prejudice found in Richeson et al. (2003) were in nearly identical areas to the regions identified in this study. The present study provides strong evidence about the functions of these regions by showing that the reduction in Black > White differences in amygdala activation in the long compared to the short duration condition was correlated with activity in dorsolateral prefrontal cortex and anterior cingulate. Thus, our findings are consistent with the idea that it is possible to control spontaneously activated negative attitudes and point to neural circuits that may be involved in this control.

These results suggest that previous neuroimaging research has failed to find robust evidence for greater amygdala activation to Black faces relative to White faces in White participants because more controlled processing can modulate more automatic processing. Greater amygdala activation to Black faces relative to White faces in the 30ms condition than in the 525ms condition is consistent with the idea that unwanted prejudicial responses are most likely to occur under conditions of distraction or cognitive overload when reflective cognitive processes that might modulate an automatically activated evaluation are otherwise engaged (Pendry & Macrae, 1999). Furthermore, Richeson et al. (2003) found that those participants who presumably engaged in more response regulation in viewing Black faces later performed more poorly on a task that required cognitive control (i.e., the Stroop task; see also Richeson & Shelton, 2003).

Such mental distraction may partially explain why a significant direct correlation between presentations of Black and White faces and the IAT was found in Phelps et al. (which used a 2sec presentation time) but not in the current study in the 525ms condition. Several methodological differences exist between these two studies. Whereas in the present study participants made left-right judgments about the location of each stimulus, in Phelps et al. attention was directed toward the faces on each trial using an n-back task. Thus, the participants in Phelps et al. engaged in intentional memory encoding operations and memory judgments for each item and may have been less able to regulate their evaluative responses. Furthermore, Phelps et al. used a blocked design in which multiple Black or White faces were presented sequentially compared with the event-related design used here in which Black and White faces were presented randomly. If control of prejudiced responses usurps cognitive resources (e.g., Richeson & Shelton, 2003), then the repeated Black faces in a blocked design may be more demanding than a design in which Black and White faces are randomly interspersed. Because the prefrontal cortex was not scanned in Phelps et al., nor were participants' motivations to control racial attitudes measured after scanning, the possibility that the Phelps et al. procedure was more reflectively demanding and may have reduced participants' control of their responses to the black faces remains a speculation.[4]

One potential contributor to amygdala activity is the ease with which stimuli can be processed and discriminated. Golby, Gabrieli, Chiao, & Eberhardt (2001) showed that same-race faces activate the fusiform gyrus -- an area associated with greater perceptual expertise (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999) -- more so than other-race faces. This result suggests that same-race faces may have an advantage as early in the processing stream as visual areas. Consistent with this idea, we found in the 30ms condition that participants' IAT

scores significantly correlated with White>Black bilateral fusiform activation. That is, the more

that participants showed a greater association between White and Good (and Black and Bad)

than White and Bad (and Black and Good), the more White faces activated the fusiform

compared to Black faces. Additionally in the 30ms condition, Black>White amygdala activity

correlated with left fusiform activity.[5] Thus, for White participants with greater race biases,

Black faces may be processed more superficially than White faces; this difference in perceptual

processing may result in a relatively undifferentiated early visual signal that triggers an

emotional response (e.g., Johnson & Multhaup, 1992).

Together, the present data provide new evidence about the neural correlates of automatic

and more controlled processing of stimuli that have conflicting valance. They show neural

differences between more automatic and more controlled processing of social groups, and

suggest that reflective processes may interact with and modulate evaluations arising more

automatically during perceptual processing.

References

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski, (Eds). *Social psychology: Handbook of basic principles*. (pp. 169-183). New York: Guilford Press.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality & Social Psychology, 62*, 893-912.

Beauregard, M., Levesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *Journal of Neuroscience. 21*, RC165: 1.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 103*, 624-652.

Cahill, L., Haier, R. J., White, N. S., Fallon, J., Kilpatrick, L., Lawrence, C., Potkin, S. G., & Alkire, M. T. (2001). Sex-related difference in amygdala activity during emotionally influenced memory storage. *Neurobiology of Learning & Memory. 75,* 1-9.

Canli, T., Desmond, J. E., Zhao, Z., & Gabrieli, J. D. (2002). Sex differences in the neural basis of emotional memories. *Proceedings of the National Academy of Sciences of the United States of America. 99,* 10789-10794.

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, I. H., & Gabrieli, J. D. (2002). Amygdala response to happy faces as a function of extraversion. *Science. 296,* 2191.

Cheesman, J. & Merikle, P.M. (1986). Distinguishing conscious from unconscious perceptual processing. *Canadian Journal of Psychology, 40*, 343- 367.

Cunningham, W. A., Nezlek, J. B., & Banaji, M. R. (in press). Implicit and explicit ethnocentrism: Revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin.*

Cunningham, W. A., Preacher, K. J. & Banaji, M. R. (2001). Implicit Attitude Measures: Consistency, Stability, & Convergent Validity. *Psychological Science, 12*, 163-170.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural Components of Social Evaluation. *Journal of Personality and Social Psychology*, *85*, 639–649.

Davis, M. & Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry, 6*, 13-34.

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

Draine, S. C. & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*, 286-303.

Eagly, A. H. & Chaiken, S. (1993). *The Psychology of Attitudes*. Harcourt Brace, Fort Worth, TX.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality & Social Psychology, 50*, 229-238.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., Frackowiak, R. S. J. (1995). Statistic parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2*, 189-210.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience, 2,* 568-573.

Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience, 4*, 845-850.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and

    stereotypes. *Psychological Review, 102*, 4-27.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual

    differences in implicit cognition: The Implicit Association Test. *Journal of Personality*

    and Social Psychology, 74, 1464-1480.

Hart, A.J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch SL (2000).

    Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli.

    NeuroReport, *11*, 2351-2355.

Johnson, M. K. & Reeder, J.A. (1997). Consciousness as meta-processing. In J.D. Cohen & J.W.

    Schooler (Eds.), Scientific approaches to consciousness (pp. 261-293). Mahwah, NJ:

    Earlbaum.

Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time course studies of reality monitoring

    and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition,*

    20, 1409-1419.

Johnson, M. K. & Multhaup, K. S. (1992). Emotion and MEM. In S. Christianson (Ed). *The*

    handbook of emotion and memory: Research and theory. (pp. 33-66). Lawrence Erlbaum

    Associates: Hillsdale, NJ.

Isenberg, N., Silbersweig, D., Engelien, A., Emmerich, S., Malavade, K., Beattie, B., Leon, A.

    C., & Stern, E. (1999). Linguistic threat activates the human amygdala. *Proceedings of*

    the National Academy of Sciences of the United States of America, 96, 10456-10459,

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human

    amygdala activation during conditioned fear acquisition and extinction: a mixed-trial

    fMRI study. *Neuron, 20*, 937-945.

LeDoux, J. E. (1996). *The emotional brain: the mysterious underpinnings of emotional life.* Simon & Schuster: New York.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*. *84*, 231-259.

MacDonald A. W., Cohen J. D., Stenger V. A., & Carter C. S. (2000). Dissociating the role of dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science*. *288*, 1835-1837.

Marcel, J.J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*, 197-237.

McConahay, J. B. (1986). Modern Racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91-125). Academic Press, Orlando, FL.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383*, 812-815.

Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and Unconscious emotional learning in the human amygdala. *Nature, 393*, 467-470.

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Precosncious control of stereotype activation through chronic egalitarian goal. *Journal of Personality and Social Psychology*, *77*, 176-184.

Murphy, S. T., Monahan, J. L., & Zajonc, R. B. (1995). Additivity of nonconscious affect: Combined effects of priming and exposure. *Journal of Personality & Social Psychology, 69,* 589-602.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with suboptimal and optimal stimulus. *Journal of Personality and Social Psychology, 64*, 723-739.

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226-254.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics, 6*, 101-115.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, 14, 1215-1229.

Pendry, L. F., & Macrae, C. N. (1999). Cognitive load and person memory: The role of perceived group variability. *European Journal of Social Psychology, 29*, 925-942.

Phelps, E. A., O'Connor K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729-738.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811-832.

Poldrack, R. (nd). downloaded from http://spm-toolbox.sourceforge.net.

Raye, C.L., Johnson, M.K., Mitchell, K.J., Reeder, J.A. & Greene, E.J. (2002). Neuroimaging a single thought: Dorsolateral PFC activity associated with refreshing just-activated information. *NeuroImage, 15*, 447-453.

Richeson, J. A., Baird, A.A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S. & Shelton, J. N. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience*.

Richeson, J.A. & Shelton, J.N. (2003). When prejudice does not pay. Effects of interracial contact on executive function. *Psychological Science,* 14, 287–290.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437-465.

Stark, C. E. L. & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences USA*, 98, 12760-12766.

von Hippel, W., & Hawkins, C. (1994). Stimulus exposure time and perceptual memory. *Perception & Psychophysics, 56*, 525-535.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*, 411-418.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72*, 262-274.

Author Note

Footnotes

---

[1] All trials, whether a face trial or white square trial, included both the forward and backward masks. Thus, because short and long white square trials were randomly presented between face stimuli, face trials in the various conditions occur within a consistent background task. The intervals between faces included these active-task filler trials to decrease uncontrolled (and presumably more heterogeneous) activity among participants during intervals between critical trials (e.g., Raye, Johnson, Mitchell et al., 2002; Stark & Squire, 2001).

[2] Because short duration trials always have a longer presentation of the abstract figure than long duration trials, we computed our comparisons of Black and White faces within each of the duration conditions. In comparisons within the same duration, the only difference between trial types is the presentation of Black and White faces. For comparisons between short and long duration trials, we computed the interaction term in an ANOVA analysis; that is, we compared the difference between activation to Black and White faces at the two levels of the duration variable. This interaction would be unlikely to be influenced by differences in brain activation from the longer presentation of the abstract figure in the short duration condition (i.e., by an effect of presentation duration).

[3] The IAT was selected to measure the implicit association between Black and good vs. bad for several reasons. First, the IAT is among the most statistically reliable implicit measures available (see Cunningham, Preacher, & Banaji, 2001). Given the small sample size of fMRI studies, a least a moderate degree of reliability is necessary to detect relationships. Second, previous work has indicated that amygdala activation to Black faces relative to White faces is correlated with IAT scores (Phelps et al., 2000), indicating that the IAT may be a valid predictor of emotional

processing. As new reliable implicit measures are developed, it will be important to additionally include such measures in future research.

[4] Another difference between Phelps et al. and the current study is that Phelps et al. found bilateral amygdala activation correlated with IAT scores, whereas in this study we found right amygdala. It is unclear what might account for such a difference. Findings to date suggest that amygdala laterality effects can vary as a function of gender (Cahill, Haier, R. J., White et al., 2001; Canli, Desmond, Zhao, & Gabrieli, 2002) or differences in personality variables such as extroversion (Canli, Sivers, Whitfield, Gotlib & Gabrieli, 2002). As more research is conducted, the reasons for laterality effects should become clearer.

[5] Relationship with IAT: Right fusiform: $t(12) = 4.26$, $p < .001$, MNI coordinates 51, -42, -24; Left fusiform: $t(12) = 4.21$, $p < .001$, MNI coordinates -45, -42, -21). Relationship with amygdala activity: Left fusiform: $t(12) = 4.66$, $p < .001$, MNI coordinates -30, -36, -21.

Table 1:

Behavioral Attitude Measures

	Mean	STD	Minimum	Maximum
IAT (Black$_{Good}$ – Black$_{Bad}$)RT	160.17	154.14	-101	465
Combined Attitude Index	1.35	0.28	0.85	1.79
Modern Racism	2.15	0.31	1.71	2.57
Internal Motivation	4.82	0.51	4.20	6.00
Implicit-Explicit Conflict	4.88	5.13	-3.48	15.57

IAT = Implicit Association Test; Modern Racism = Modern Racism Scale; Internal Motivation = The Internal Subscale of the Motivation to Respond without Prejudice Scale. The Combined Attitude Index was computed by averaging the Internal Motivation Scale with the Modern Racism Scale (reversed scores) and then subtracted the scale midpoint (3.5).

Table 2:

Areas of significant BOLD activation

Black>White (30ms)

					MNI coordinates		
Size	**Area**	**BA**	**R/L**	**t**	**x**	**y**	**z**
11	Amygdala	n/a	R	4.26	18	-6	-12
32	Superior Frontal Sulcus	9	L	4.79	-21	30	39
27	Superior Temporal Cortex	36	R	4.88	30	-6	-33
27	Supplementary Motor	6	L	4.56	-3	0	66
		6	L	4.05	-48	-6	54

Black>White (525ms)

77	Anterior Cingulate	32	L	5.82	-6	36	24
		32	L	4.52	-9	18	33
		32	R	4.34	3	18	33
13	Ventrolateral PFC	47	R	4.04	57	30	-12
37	Dorsolateral PFC	9	R	4.88	27	48	24
		9	R	4.85	33	48	36
15	Dorsolateral PFC	10	R	5.22	24	60	27
20	Supplementary Motor	6	LR	4.12	0	0	72

White>Black (30ms)
 None

White > Black (525ms)

23	Dorsolateral PFC	44	L	5.64	-57	18	36
16	Hippocampus		R	4.91	33	-24	-18

Tables show local maxima > 8.00mm apart per cluster. x, y, and z coordinates are in MNI (Montreal Neurological Institute) space. BA = Brodmann's Areas. R/L = Right or Left hemisphere.

Figure Captions

Figure 1: Activation in right amygdala: 30ms Black– White faces contrast

(A) Coronal random effects contrast map showing activation p < .005 for the 30ms Black>White

contrast, MNI x,y,z coordinates 18, -6, -12. (B) Timelines for both the 30ms and 525ms Black

and White faces were generated from these significant voxels. Because participants were

presented with abstract figures and white squares between face presentations, these time courses

should be thought of as relative to these filler trials.

Figure 2: Prefrontal activations: 525ms Black – White faces contrast

Statistical maps were generated from the 525ms Black>White contrast, p < .005. Contrast maps

for significant voxels are presented separately for the (A) dorsolateral prefrontal cortex (MNI:

57, 30, -12; 27, 48, 24; 24, 60, 27); (B) ventrolateral prefrontal cortex (MNI: 57, 30, -12); (C)

anterior cingulate (MNI: -6, 36, 24; 3, 18, 33). For larger areas of activation, multiple MNI x,y,z

coordinates are presented for areas of local maxima greater than 8mm apart. Data are presented

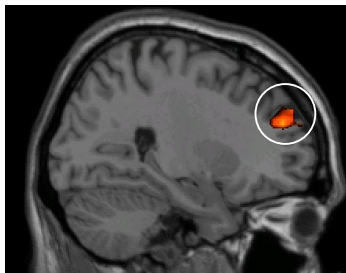in sagital orientation to show the extent of activation.
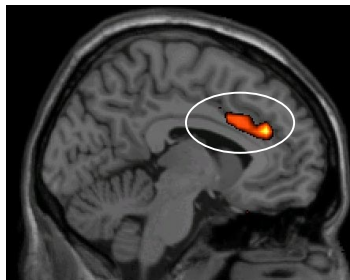
(A) Amygdala contrast map

(B) Amygdala timecourses
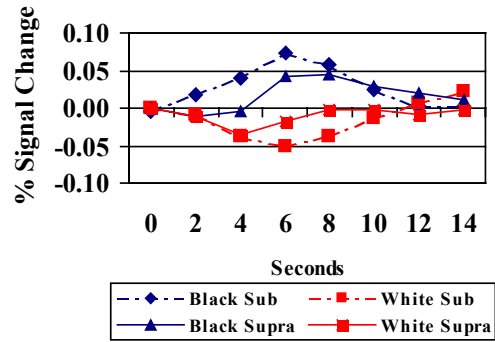
(A): dorsolateral PFC　**(B): anterior cingulate**　**(C): ventrolateral PFC**