The Pennsylvania State University

The Graduate School

College of the Liberal Arts

READING BETWEEN THE LINES OF THE FACE:
THE RELATIONSHIP BETWEEN MENTALIZING AND FACE MEMORY

A Thesis in

Psychology

by

Robert G. Franklin, Jr.

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2009

The thesis of Robert G. Franklin, Jr. was reviewed and approved* by the following:

Reginald B. Adams, Jr.
Assistant Professor of Psychology
Thesis Advisor

Theresa K. Vescio
Associate Professor of Psychology

Michael J. Wenger
Associate Professor of Psychology

Susan Mohammed
Associated Professor of Psychology
Graduate Program Chair

*Signatures are on file in the Graduate School.

ABSTRACT

Of the many meanings communicated by a face, decoding what a person is thinking and remembering a person's facial identity are two essential processes for smooth social interactions. This thesis examines the relationship between decoding a person's mental state based on facial cues and the ability to remember that person's face. To examine this relationship, I first tested whether differences in the extent to which faces elicit mental state reasoning are associated with corresponding improvements in face memory. I predicted that the more a face elicited more complex mental state reasoning, the better it would be remembered. Additionally, I predicted that the better participants were at decoding mental states from others, the better they would also remember new faces in a separate memory test. Study 1 supported both of these hypotheses. Secondly, I examined whether face memory mediates the common ingroup bias found in face memory. Ingroup bias is defined as better memory for faces one perceives as belonging to their own group. I observed mixed evidence in support of this. Surprisingly, this study failed to replicate an ingroup bias in face memory using the surprise memory test. However, when I used a memory test in which participants were explicitly told to remember faces for a subsequent memory test, I then replicated the own-group bias and found a trend toward mentalizing as a mediator of this relationship. This incidental finding suggests that the own-race bias is in some ways a distinct phenomenon compared to the more general own-group bias, and explanations for how this may be the case are discussed.

TABLE OF CONTENTS

Chapter 1: Reading between the Lines of the Face: The Relationship between Mentalizing and Face Memory

The face holds a unique role in social communication. In the contours and structure of the face, many different forms of information relevant to social behavior exist. We can decode identity, age, gender, race, mental state, and behavioral intent in an instant, with little more than a glance. Some researchers have speculated that the faces are somehow “special” in the way they are processed, claiming dedicated brain regions are specialized for face processing (Farah, Wilson, Drain, & Tanaka, 1998; Kanwisher, McDermott, & Chun, 1997; but see Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Even though debate exists as to whether there exists dedicated neural underpinnings specialized for face processing, no debate exists that faces serve a special role in social communication. No other stimulus conveys such a large amount of socially relevant information in such a small and overlapping array of features as the face. Given the unique place the human face holds in social communication and person construal, it stands to reason that the ability to remember faces may be influenced by the extent to which social meaning is communicated by them. This thesis investigates the hypothesis that the degree to which one reads social meaning into a face to determine what another is thinking or feeling influences how well faces are remembered. This hypothesis draws on and has implications for both social psychology and visual cognition. Therefore, it is necessary to review and connect these otherwise disparate literatures in building the foundation of this thesis.

Given that this thesis represents a nexus of different fields, I first wish to describe and define the various terms that will be used herein. Many researchers have examined the construct of *mentalizing*, which refers to a person's ability to decode and understand the mental states of others (Amodio & Frith, 2006; Castelli, Happe, Frith, & Frith, 2005; Frith & Frith, 2001).

Though these literatures do not describe mentalizing as specific to face perception per se, convincing evidence, reviewed below, asserts that humans are able to understand another's mental state using just facial cues. Therefore, in this thesis, the term mentalizing will be used to describe the process by which we infer another’s mental states based on facial cues. Related to this construct, I will argue that social factors can affect the degree to which people mentalize about a face and thus the level of complex social meaning that is read into a face. Different situational factors and relational differences between people may moderate how motivated they are to process what it is another is thinking, and thus these factors may influence the extent to which they mentalize about faces. Likewise, I also hypothesize that certain facial cues themselves may vary in the amount of social meaning they convey. As we are adept at reading personality traits and mental states from the face itself (e.g., Zebrowitz, 1997), I hypothesize that differences in stimulus features will influence the degree to which mentalizing takes place.

Social Factors and Face Memory

Group Memory Effects

The effects of group membership are present in a wide variety of social categories, from very rigidly defined group membership, such as race or gender, to seemingly trivial person groupings, such as randomly assigned workgroups. Ingroups, or groups of which a person is a member, are normally favored over outgroups, or groups of which a person is not a member (Tajfel, 1970). Merely creating minimal groups—groups that are randomly assigned and have no basis outside of an experimental setting—is all that is necessary to give rise to such effects (e.g., Sherif, 1956), underscoring the power of self-perceived group membership in social interactions. Group membership also creates biases in how individuals are perceived. People tend to attribute less individual characteristics (i.e., decreased individuation) to outgroups relative to ingroups, a concept referred to as outgroup homogeneity (Judd & Park, 1988; Ostrom & Sedikides, 1992). Outgroup homogeneity generalizes to both social groups traditionally present within a culture and to minimal arbitrarily assigned groups. Essentially, outgroup homogeneity can been summarized as: “we are all unique but they are all the same.”

Many studies of group perception and outgroup homogeneity have used face identification and face memory paradigms to study this effect on a perceptual level (Boldry, Gaertner, & Quinn, 2007). The most extensively researched of these is the own-race bias. The own-race bias is an effect where same-race faces are consistently remembered better than other-race faces (for reviews, see Meissner & Brigham, 2001; Sporer, 2001). The own-race bias in facial memory is most often studied in variations of identification tasks where participants are shown a series of faces of one's own race and another race and then are asked to discriminate between these faces and novel ones (e.g., Malpass & Kravitz, 1969). The own-race bias generalizes to other memory tasks as well, including police photo lineups (Brigham & Ready, 1985) and reaction time studies of face recognition (Valentine, 1991), thereby demonstrating its validity and relevance in several different forms of memory paradigms and also in real-world applications.

Even though the own-race bias is by far the most studied group effect in memory, membership in other social groups also affects face memory. Women perform better overall in face memory tasks than men (Shapiro & Penrod, 1986), arguably because they have better memory for female faces, demonstrating an ingroup advantage (Lewin & Herlitz, 2002). Some studies have reported a corresponding advantage for men when remembering male faces (Wright & Sladden, 2003), though less empirical support exists to confirm this advantage (e.g., Rehnman & Herlitz, 2006). Own age biases also exist as faces of people who are approximately the same age as a perceiver are remembered better than faces of other age groups (Wright & Stroud, 2002). Biases for ingroup faces extend to arbitrarily assigned groups as well. Participants show better memory for members of their own group than another group in a traditional minimal group paradigm, where faces are randomly assigned to be either members of one's university or a rival university (Bernstein, Young, & Hugenberg, 2007).

Moderators of Face Memory

Since several different ingroups show memory biases, a parsimonious explanation for these biases would assert that the same mechanisms cause these effects. Given this, several different explanations exist to explain why faces of one's own group are remembered better than other groups. Though most of these explanations were proposed to explain the own-race bias itself, some of these can be extended to explain other ingroup biases.

Perceptual Experience and Face-space. One of the earliest explanations for the own-race bias suggested that lack of previous experience with faces of different races gives rise to same-race advantages. People tend to have less experience with other-race faces than faces of their own race, and it is possible that this lack of experience hindered the encoding of more nuanced differences between other-race faces. Support for this idea came from the experiences of racial minorities, who often showed either an attenuated own-race bias or no bias at all (Lindsay, Jack, & Christian, 1991; Malpass & Kravitz, 1969). People also show larger biases in other races with which they have had little contact than races with which they are more familiar (Chance et al., 1975; Sporer, 2001).

The perceptual model of facial encoding in multidimensional face-space is a cognitive model that can be extended to explain why perceptual experience may affect face memory (Valentine, 1991). The face-space model posits that each face is encoded as a node in multidimensional space. In this model, each dimension represents a different useful trait for discriminating faces, such as a difference in spacing between facial features. One way that faces can be represented in this framework is norm-based encoding. In norm-based encoding, facial identity is remembered based on how far each dimension differs from a prototypical or average face, which represents a zero point for each dimension (Leopold, O'Toole, Vetter, & Blanz, 2001). In this model, norms represent dimensions in the multidimensional space framework and the face is encoded on each norm by virtue of how much it differs from the prototypical face. Supporting the norm-based network, caricatures of faces, which are faces modified to deviate more from average, are easier to identify than anti-caricatures, which are faces that are modified to more approximate the average (Lee, Byatt, & Rhodes, 2000). Additionally, when the perceptual average is moved away from a face by using face adaptation aftereffects, faces are easier to recognize (Leopold et al., 2001; Leopold, Rhodes, Müller, & Jeffery, 2005).

Norm-based encoding can be used to explain own-race biases in memory. Some norms exist that are specific for coding faces of only one gender or one race (e.g., Little, DeBruine, & Jones, 2005). If this is the case, then it is possible that more norms exist for encoding faces of one's same race or gender, which makes it easier to identify and remember those faces within the norm-based encoding model. Recent visual aftereffects support the assertion that some separate norms exist for faces of separate races and genders (Jaquet, Rhodes, & Hayward, 2007). In this case, repeatedly showing faces of one race caused visual adaptation aftereffects most strongly for faces of that race specifically and were reduced when generalized to other faces. Since some norms are specific to race, perceptual experience and encoding of faces with fewer norms could explain other-race deficits in face memory.

Even though perceptual expertise and the underlying mechanism suggested by norm-based encoding of race may explain some of the own-race bias, it is not sufficient to explain the own-race bias (Meissner & Brigham, 2001). It is also not adequate in explaining other own-group biases. Ostensibly, people have ample experience in seeing faces of both genders, yet there exists some evidence for an own-gender bias. Even more compelling evidence against perceptual expertise is that faces randomly assigned to be members of an ingroup are remembered better than faces randomly assigned to an outgroup, which cannot be explained by any perceptual expertise toward those faces (Bernstein et al., 2007; see also MacLin & Malpass, 2001). This indicates that differences in encoding faces of different races based on experience with those faces itself are not sufficient to explain the own-race bias.

Social Categorization. Attitudes about one's group versus other groups may drive face memory effects. Ingroup face processing biases may be related to more global differences in social processing, specifically those involving how socially relevant a stimulus is (Adams, Pauker, & Ambady, 2009). People tend to associate more secondary or complex emotions to ingroup members and more primary or basic emotions to outgroup members (Paladino et al., 2002). Ascribing complex emotions, which are considered uniquely human mental states, to only certain groups of humans is referred to as infrahumanization (Leyens et al., 2003). Infrahumanization is present in a variety of group settings, increasing when groups are stereotyped and decreasing when group members are individuated, supporting the assertion that group membership per se can affect how others attribute basic or complex mental states to another (Cuddy, Rock, & Norton, 2007).

The attitudes that a person has about another and how a person categorizes that person based on a person's group membership can interact to influence face memory as well. For example when a Black face, which is stereotypically associated with anger (Hugenberg & Bodenhausen, 2003), is shown displaying a threatening emotion, memory accuracy increases (but see Corneille, Hugenberg, & Potter, 2007). This is ostensibly because an outgroup member with a threatening face is more functionally relevant to remember than an outgroup member with a nonthreatening expression (Ackerman et al., 2006). Adding stereotypical information that is associated with a group can itself cause memory differences for faces of different groups. For instance, merely changing a face's hairstyle from a same-race stereotypical hairstyle to an other-race stereotypical hairstyle can cause an own-race bias (MacLin & Malpass, 2001). Likewise, adding gender-stereotypical hairstyles increases the degree to which same-gender faces are remembered independent of the effect of the gender of the face itself (Wright & Sladden, 2003). In this case, how stereotypical a hairstyle was rated additionally contributed to the memory effect above the effect of the face gender itself. This is also the case with hairstyles that are stereotypically associated with heterosexual versus homosexual males, which also gives rise to cross-group memory effects (Rule, Ambady, Adams, & Macrae, 2007). This work indicates that categorizing faces based on group membership itself yields memory effects for faces, even in faces that are otherwise identical.

To explain ingroup advantages in face memory, Levin (1996, 2000) suggested that the own-race bias is caused by differences in the application of social cognitive categories to ingroup and outgroup members. Categorizing faces as outgroup members leads to only surface level processing of those faces, whereas categorizing faces as ingroup members leads to deeper-level processing. This distinction in visual processing is consistent with the concepts of outgroup homogeneity and individuation describe above. In this case, individuation reflects judgments that are based on processing that go beyond the level of a mere categorical feature (Levin, 2000).

Such a model accounts for the effects of experience in the own-race bias and can explain deficits for any outgroup, even arbitrarily assigned groups. Individuation results in an additional deeper level of processing (e.g. Craik & Lockhart, 1972), whereas processing an outgroup face presumably results in lowered motivation to process faces deeply (e.g. Hugenberg & Sacco, 2008). What this additional level of processing entails is not clear, but it may involve increased encoding within the face-space framework discussed above.

In support of the social categorical explanation, group biases in memory extend to more ambiguous social categories with memory varying as a function of same- versus other-group categorization, including ambiguous race races (Pauker, Weisbuch, Ambady, Sommers, Adams, & Ivcevic, *in press*), and sexual orientation (Rule, Ambady, Adams, & Macrae, 2007). In cross-race memory, group biases only occur in direct relative to averted gaze faces, where social categorization and social relevance is enhanced (Adams et al., 2009). Further, faces that are even arbitrarily labeled as students of an outgroup college or different socioeconomic class are remembered less accurately than faces labeled as part of one's own college or class (Bernstein, Young, & Hugenberg, 2007; Shriver, Young, Hugenberg, Bernstein, & Lanter, 2008). Taken together, these findings suggest that ingroup biases cannot be due merely to systematic variation in different group faces or due to lack of exposure to these faces, but must be due to a more generalized cognitive mechanism (e.g., Eberhardt, 2006; Hugenberg & Sacco, 2008). In all of these cases, the suggested mechanism is that outgroup faces are processed less deeply because they are categorized while ingroup faces are individuated and thus processed deeper.

Additionally, work on infrahumanization suggests individuation includes the tendency to ascribe more complex mental states to ingroup than outgroup members (Paladino et al., 2002; Leyens et al., 2003). Supporting this contention, Adams et al. (in press) demonstrated that

Japanese and U.S. Caucasian participants performed better when decoding complex mental states from eyes of same-culture versus other-culture faces. This was associated with greater neural activation in the superior temporal sulcus for ingroup faces versus outgroup faces, suggesting that deeper mental processing may be associated with performance in decoding mental states.

Current Research

The research reviewed above describes how group dynamics affect the nature of face memory as well as interpreting mental states of others. However, the act of inferring mental states from others, also commonly referred to as mentalizing (Amodio & Frith, 2006; Brothers, 1990; Castelli et al., 2005; Frith & Frith, 2001), has not previously been incorporated in models of face perception or memory (Adams, Franklin, Nelson, & Stevenson, *in press*; Franklin & Adams, under review –c; Zebrowitz, 2006). In fact, current models of face perception assume that the processing of expressive cues, such as those that indicate one's mental state, and the processing of identity are distinct processes with dissociable neural underpinnings (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000). Given that these models of face processing treat identity and expression as doubly dissociable processes, it is perhaps not surprising that the role of mentalizing has not previously been examined in relation to face memory.

Here I discuss three factors which may influence mentalizing faces which is in turn predicted to influence face memory (see Figure 1). *Stimulus factors* are those factors within a face itself that may influence the degree to which complex mental states are read into a face, and thus how memorable it is. To my knowledge, no previous research has documented systematic variation in mentalizing due to facial appearance, let alone establishing a link between mentalizing and face memory. Next, *perceiver factors* represent individual differences such as empathy and nonverbal sensitivity that may influence mentalizing and thus face memory. Again, to my knowledge a link between nonverbal indicators of mental state decoding and face memory has yet to be established. Finally, *relational factors* are those defined by the relation of a perceiver to a stimulus person that may influence mentalization and face memory, such as own-group biases as discuss above. Again, although evidence exists for cross-group effects in both memory and mental state decoding, to my knowledge the two have not been related to one another. In attempting to establish the relationship between mentalizing and memory, the current thesis aimed to examine the relationship between mentalizing and face memory along all three of these factors, each of which is described in more detail below.

Stimulus Factors

Within a group of faces of the same age, gender, or race, I predict that individual differences will still exist in the extent to which certain faces inspire mentalizing. Faces have variable affective and reward value based on how attractive they are (Aharon et al., 2001, Franklin & Adams, under review-a; -b; O'Doherty et al., 2003). Additionally, person perception from faces systematically varies based on facial features that give rise to perceived babyishness versus maturity (Berry & McArthur, 1985; Zebrowitz, 1997), as well as dominance and sociability (Knutson, 1996; Hess, Blairy, & Kleck, 2000, Mignault & Chaudhuri, 2003). These are just a few examples of how stimulus factors can affect how faces are socially perceived.

Stimulus factors also affect how memorable faces are. For instance, both attractive and unattractive faces are more memorable than faces of average attractiveness (Shepherd & Ellis, 1973). Faces with direct eye gaze are better recalled than those showing averted eye gaze (Mason, Hood, & Macrae, 2004) and emotional expression can interact with eye gaze to influence face memory (Franklin & Adams, under review-c). Thus, attractiveness, emotional expression, and eye gaze are all factors that affect memory by changing the relevance the faces hold to a perceiver. However, the extent to which a face inspires mentalizing arguably goes beyond these aspects of face perception. Herein, I argue that some faces have a quality that makes them more socially engaging to a perceiver and thus elicit more mentalizing. I contend that there may exist features within the face that give rise to differences in the extent to which observers tend to mentalize about the face, making some faces seem to convey more complex mental states and others seem vacuous. In turn I predict that the extent to which a face elicits mentalizing will influence the depth to which a face is processed, and thus remembered, as deeper levels of processing are associated with greater memory (Craik & Lockhart, 1972). It is beyond the scope of this thesis to determine what exactly the physical qualities are that inspire such engagement in the first place. That such a quality exists in the face or that such a quality is related to face memory needs first to be established. Thus the first aim of this thesis is to establish that faces do vary systematically in the extent to which they elicit complex mental states. I predict that there will be consensus across observers in the extent to which faces convey mental state complexity and that this will in turn predict what faces are better remembered.

Perceiver Factors

The above section predicts that there will be individual differences in the extent to which certain faces trigger the process of mentalizing in faces. In this section I assert that stable individual differences in perceivers' abilities to infer a person's mental state from nonverbal information, or nonverbal sensitivity, will also predict face memory. I predict that individuals who perform better in tasks that measure nonverbal sensitivity based on reading emotions and mental states from faces will also be better at remembering facial identity based on tests employing separate invariant facial appearance cues. Performance in tasks of nonverbal sensitivity based on facial cues is known to vary widely across individuals (e.g. Baron-Cohen et al., 2001; Depaulo & Friedman, 1998). For instance, women perform better at mental state decoding tasks than men (Baron-Cohen et al., 2001; Hall, 1978; Zuckerman, Lipets, Koivumaki, & Rosenthal, 1975).

Several different cognitive and affective processes are responsible for the differences people exhibit in nonverbal sensitivity. The construct of attentional accuracy, or the degree of attention paid to nonverbal cues, is critical to understanding nonverbal cues and nonverbal sensitivity. Those who have higher attentional accuracy also show better memory for nonverbal cues, and thus have a higher level of nonverbal sensitivity (Hall, Carter, & Horgan, 2001). Part of this advantage in nonverbal sensitivity is related to the ability to more efficiently process information and multitask in social interactions (Lieberman & Rosenthal, 2001). Therefore, it is reasonable to suppose that nonverbal sensitivity can play a role in face memory, yet to my knowledge no research has examined this relationship. Nonverbal sensitivity may represent a way to measure the construct of individual differences in the extent to which one mentalizes from faces.

Relational Factors

Relational factors are intergroup variables defined by an interaction between stimulus and observer. Own-group biases are an example of the importance of relational factors in face memory. Relational factors can also be influenced directly by certain perceiver factors. The influence of how one perceives his or her own identity in relation to another reflects a process by which a face can take on special meaning only in relation to the perceiver. So, where both a perceiver and a stimulus face have a race associated to them, whether a perceiver sees a face as a same-race face or an other-race face depends on the relationship between the race associated with a perceiver and the race associated with the stimulus. As already reviewed above, whether a face is of an ingroup or outgroup member can exert a strong influence on face memory. Differences in group memberships also moderate differences in mental state decoding. In addition, people tend to perform better at mental state decoding when presented with same- versus other-culture faces (Adams, Rule, et al., *in press*). Given that people appear more motivated to individuate own-group versus other-group faces (Sacco & Hugenberg, 2008), I also hypothesize that differences in the extent to which people mentalize about ingroup versus outgroup faces will mediate the influence of group membership on face memory.

Chapter 2: Method

Experiment 1

Contemporary models of face processing assume that the processes involved in decoding identity and decoding another's mental state are separate and dissociable. Therefore, little research has examined if there exists a relationship between mental and emotional state decoding and face memory. Experiment 1 examined this relationship by testing the role of perceiver and stimulus factors on this relationship. This study examined individual differences in observers in their ability to read mental states from others and individual differences in stimulus faces that elicit greater amounts of mental state decoding and how each of these were related to face memory. This study was divided into two phases. Phase 1 tested perceiver factors by examining if those people who show higher levels of nonverbal sensitivity are also better at remembering faces. Phase 2 examined the role of stimulus factors in memory by testing if those faces that evoke more mentalization are better remembered.

Phase 1: The Influence of Individual Differences in Observers' Mental State Decoding Ability

To test the first question, the relationship between individual differences in observers' mental state decoding ability and memory, I used two separate tests of nonverbal sensitivity. The first test was the Reading the Mind in the Eyes Test (Eyes Test, Baron-Cohen et al., 2001). This test consists of showing the eye region of 36 faces and having participants choose between one of four complex mental state adjectives to describe the eyes. The second test was the Diagnostic Analysis of Nonverbal Accuracy (DANVA, Nowicki & Duke, 2001). The DANVA consists of 24 pictures of faces showing what are described as high or low intensity basic emotion states where participants have to choose if a face is displaying an angry, fearful, happy, or sad expression. The Eyes Test represents a well-validated test of complex mental state decoding while the DANVA tests discriminations of high and low intensity basic emotions (i.e., anger, fear, joy, and sadness).

Participants. Sixty-one undergraduate students (30 female, 31 male) completed the study for course credit.

Stimuli. Twenty-four female and 24 male grayscale images of Caucasian faces were used. The faces were cropped within an ovoid region to include the face and to exclude hair cues, due to the fact that hair cues are particularly important in face recognition (O'Donnell & Bruce, 2001). Cropping hair cues is not always done in face memory studies; however, because hairstyle can have considerable influence in top-down modulated social categorization (cf. MacLin & Malpass, 2001; Rule et al., 2007), it was necessary to establish that these effects are specific to the face itself by cropping out hairstyles. The faces were randomly divided into two sets of 24 faces, balanced for gender. For each participant, one set served as the target stimuli while the other served as the distracter stimuli. The set that served as the target or distracter stimuli was randomly counterbalanced across participants so that an approximately equal number of participants saw each block.

Two measures of nonverbal sensitivity were also employed in the design. The first of these was the Eyes Test. This task consists of presenting the eye regions of 36 faces with four descriptive adjectives. In the Eyes Test, photographs were taken from magazines and cropped to only expose the eye regions of faces. Target words and foils were chosen by two of the original authors and pilot tested to create a consensus response for which word best described each face (see Baron-Cohen et al., 2001). During the test, participants chose which of these adjectives best describes the mental state of the person. The Eyes Test is well-validated and has convergent validity with traditional tests of mentalizing. This test has been administered in over sixty studies, examining both clinical and nonclinical populations. Though the Eyes Test was developed to examine differences between normal populations and Autism Spectrum Disorders, it has been found to be predictive of a variety of different clinical disorders marked by impaired social perception, as well as intergroup differences (Adams et al., in press).

The second test used was the Diagnostic Analysis of Nonverbal Accuracy (DANVA, Nowicki & Duke, 2001) adult facial expression scale. The DANVA uses pictures of adult anger, fear, joy, and sadness expressions which vary in intensity. Twelve of the faces are described by the authors of the test as high-intensity expressions while 12 others are low-intensity expressions and thus harder to decode. The DANVA is a well-validated test of the ability to recognize basic emotions used in a variety of clinical samples, especially in studying developmental disorders (Nowicki & Carton, 1993). The DANVA correlates with other measures of nonverbal sensitivity, but may represent a more generalized process of nonverbal sensitivity (Rosip & Hall, 2004).

Method. This experiment employed a memory paradigm followed by the two tests of nonverbal sensitivity. The memory phase was divided into an encoding and recall block with a distracter task in between. Prior to beginning the paradigm, the experimenter explained to participants that they would see a series of faces and that they were to try to remember those faces for a subsequent recall task. Thus, this was an intentional memory paradigm since participants were instructed to explicitly remember the faces. A distracter task immediately followed the encoding task during which participants completed math problems before the recall task was administered. Written instructions appeared on the screen to remind participants this during each phase of the experiment.

During the encoding task, 24 target faces appeared on the screen for three seconds in a random order, each separated by a 500 ms fixation cross. After the encoding task, participants completed a brief four minute distracter task which consisted of simple arithmetic problems (e.g., 12 – 5, 6 x 2). Each problem was shown on screen for ten seconds during which the participant answered by typing the answer in the computer with the keyboard. The experiment continued at a consistent pace for each participant in order to ensure that an equal amount of time separated the encoding and recall tasks. The math problem task was preceded by a 15 second instruction screen that explained the math problem task to the participants and following this task was a 15 second instruction screen that reminded participants that they would be engaging in a recall task, totaling 4 ½ minutes. The recall task displayed the 24 target images along with the other 24 images as distracter images. Participants responded with the mouse if they had or had not previously seen the face before. Each face was displayed for five seconds, during which the participant responded. Following this memory task procedure, participants completed the DANVA and the Eyes Test, with the order of these measures counterbalanced across participants.

Results. Signal detection analysis (*d'*) was used to calculate recognition performance. Participants scored moderately well on the memory task (overall accuracy: *M* = .68, *SE* = .12; *d'* = 1.90, *SE* = .09), on the DANVA (*M* = 76.9%, 18.45 out of 24, *SE* = .33), and on the Eyes Test (*M* = 73.3%, 26.39 out of 36, *SE* = .51). These scores are well within the normal range for adults performing the task. DANVA scores were divided into high and low intensity emotions because high intensity emotions may have been affected by ceiling effects (*M* = 10.78 out of 12, *SE* = .17). High intensity emotions had a skewness of -1.91 (*SE* = .31) and a kurtosis of 5.59 (*SE* = .60), while low intensity emotions had a skewness of -.418 (*SE* = .31) and a kurtosis of .307 (*SE* = .60). This indicates that low intensity emotions were likely not affected by ceiling or floor effects, whereas high intensity emotions were, making any results based on high intensity emotions potentially spurious. To avoid these issues, only low intensity emotions are reported below.

Correlations were used to assess the relationship between individual differences in the mental state decoding measures and face memory. Correlations showed low intensity DANVA emotions were correlated with Eyes Test scores ($r(58) = .312$, $p < .05$). Low-intensity DANVA scores were also correlated with memory($r(58) = .261$, $p < .05$). However, Eyes Test scores were not correlated with memory, $r(58) = .08$. In order to test to see if the low-intensity DANVA scores were a significantly better predictor of face memory ability than the Eyes test, I computed a comparison of these within-subject correlations using the Z_2^{*}*bar* statistic described in Steiger (1980). This statistic is Gaussian distributed and an adjustment of the Fisher-Z statistic for within-subject comparisons of correlations. This examination showed a significant difference between the correlations for the overall DANVA scores and the RME scores, (Z_2^{*}*bar* = 2.11, $p < .05$) indicating that the DANVA was indeed the better predictor of face memory.

Gender Effects: The effects of gender were tested as well by computing a 2(gender of face) X 2(gender of participant) mixed-design ANOVA using signal detection scores. This revealed no significant effects (all Fs < 2), failing to replicate previous work that has found same-versus other-gender effects.

Phase 2: Stimulus Differences

Phase 2 examined if individual differences in neutral facial stimuli can influence the extent to which they triggered mental and emotional state decoding in observers and thus the extent to which they were remembered. Specifically, using the set of stimulus faces described above, I tested whether certain faces consistently evoke greater mental state decoding than other faces, and whether these faces are in turn better remembered. To do this, I assessed the degree to which participants rated faces on a continuous scale to gauge mental state complexity. Participants were asked to rate faces on a scale that measured the degree to which they read the mental state of a stimulus face. These ratings were then correlated with the face memory accuracies collected in Phase 1 to see if they were a predictor of memory.

Participants. Twelve female and 12 male undergraduates completed this study for class credit. They were separate participants than those in Phase 1.

Method. Raters completed continuous ratings of the same faces on their perceived mental/emotional state complexity. For the explicit mental state scale, participants were instructed that they would see a series of faces and would rate them on a seven point scale on how complex of a mental state they had. Participants were told that if the face did not appear to convey much mental state to rate the face lower on the scale, if the face appeared to convey basic emotional states, such as anger, happiness, or fear, to rate the face more toward the middle of the scale, and if the face conveyed more complex mental states, to rate the face higher on the scale. In this scale, higher numbers indicated attributing more complex mental states to faces. The seven-point scale was anchored with 1 being "no mental state," 4 being "basic emotion," and 7 indicating "complex mental state."

Results: Ratings on the continuous scale of mental state complexity were reliable across participants ($\alpha = .717$). Mental state decoding was significantly correlated with memory, ($r(46) = .338$, $p < .02$). These findings indicate that the extent to which mental state complexity is read into a face is positively related to how memorable the face is.

Discussion

Experiment 1 found evidence supporting a relationship between mental and emotional state decoding and face memory at both the perceiver and stimulus levels. Scores on the low-intensity DANVA correlated significantly with face memory, indicating that those who were better at decoding emotion low intense expressive faces were also better at remembering a separate set of neutral faces. Notably, no relationship was present between memory and the Eyes Test, which suggests that individual differences in complex mental state decoding from just the eye region of the face were unrelated to individual differences in face memory. The failure to find a correlation with the Eyes Test and face memory may have something to do with this test's focus on extracting mental states from just the eye region of the face rather than from the whole face. Future research examining complex mental state decoding from the whole face and its relation to face memory therefore warrants future investigation.

A relationship between mentalizing and memory was also found at the stimulus level. This experiment provided evidence that certain faces consistently evoked greater mentalization than others, as measured by the degree to which participants attributed more complex mental states to them. These ratings were also positively correlated to face memory, indicating that the faces that elicit greater complex mental state decoding were also the ones that were better remembered.

Experiment 2

Experiment 1 provided initial support for the proposed relationship between reading mental and emotional states in faces and face memory. A stronger test of this relationship would be evidence that mentalizing mediates cross-group memory effects. Thus, Experiment 2 aimed to examine if differences in mentalizing also mediate the own-group memory bias. Ingroup biases are present in both face memory and mental state decoding, which suggests a potential relationship may exist between the two processes particularly in light of the findings of Experiment 1. Here, I propose that the degree to which one reads mental and emotional states into an ingroup face evokes deeper processing of that face compared to an outgroup face. That in turn leads to the face being more deeply encoded and thus remembered better.

This experiment used a variation of the minimal group paradigm employed by Bernstein et al. (2007). I hypothesized that arbitrarily assigning faces to represent ingroups versus outgroups would influence the extent to which participants mentalize about these faces. I further predicted that the relationship between group membership and face memory would be mediated by the degree to which participants also differentially attributed mental states to those faces.

Participants: Ninety-four undergraduate participants completed Experiment 2 for course credit, (54 female, 41 male). Eighty-two participants were right-handed, 13 were left-handed.

Stimuli: Ninety-six stimuli of Caucasian faces were used for this study. Ovoid regions of the face area were cropped from grayscale facial images in the same way as in Experiment 1. These images were standardized for luminosity and sized so that the faces were all the same size and occupy approximately the same area of the cropped region.

For each participant, stimuli were divided into two groups, where half were presented as Penn State students, with a dark blue background and surrounded by the school’s logo to prime ingroup identity, similar to Bernstein et al., (2007). The other half were presented as students from the University of Michigan, a rival school, with a maize background surrounded by that school's logo. Faces were counterbalanced so that each face was shown to approximately the same number of participants as an ingroup target, an ingroup distracter, an outgroup target, or an outgroup distracter. Therefore, each face was shown equally often as an ingroup face or outgroup face.

Design and Procedure: Participants engaged in a task where they are asked to mentalize about the target images, completed a brief distracter task, and then completed a surprise memory task. Participants arrived to the lab in groups of two and were instructed that they would be participating in an experiment with collaborators at the University of Michigan about determining how we read what people are thinking. Participants were explicitly told that "we are working with collaborators at the University of Michigan to determine the degree to which we can read emotion into people's faces. These faces represent students, staff, and faculty from these universities." This was done in order to invoke a group manipulation by using the same type of categories that Bernstein et al. (2007) used. However, participants were told to complete a different task during the encoding phase of the experiment. Participants saw the target images in a paradigm where they were asked to describe "which one word best describes what these people are thinking or feeling at this time" and were not instructed that they would have to remember the faces in an upcoming memory test. Faces were displayed for three seconds after which a fixation screen appeared and participants responded by typing one word into a text box presented on the computer.

After doing a brief 2 minute distracter task involving simple math problems, the participants performed a surprise memory task where they were asked if they have seen the faces that were shown in the mentalizing task. Participants used the keyboard to respond to the faces with the "J" key (right index finger) to indicate a face that they had previously seen and the "F" key (left index finger) to indicate a novel face. The keys were reversed for left-handed individuals in order to pair the dominant hand index finger with a hit response. Following this, participants completed the two measures of nonverbal sensitivity used in Experiment 1.

Results

Subject-Level Analysis: Mean overall accuracy was 65.4 % (*SD* = .086). Recognition performance was computed on a subject level by calculating signal detection measures (*d′*). Replicating the findings in Experiment 1, recognition performance correlated with performance on the DANVA low-intensity emotions ($r = .237$, $p = .028$). As before, memory scores were not correlated with Eyes Test, though there was a near trend toward a correlation ($r = .157$, $p = .143$). In this case, there was also no significant difference between the correlation of the DANVA and face memory and between the Eyes Test and face memory, $Z_2^{*}bar = 1.37$, $p = .17$. Therefore, in this sample, I cannot conclude that the DANVA is significantly more predictive of face memory than the Eyes Test.

The effects of gender were tested as in Experiment 1 by computing a 2(gender of face) X 2(gender of participant) mixed-design ANOVA using signal detection scores of each individual subject. This revealed main effects of gender of face ($F(1, 93) = 26.15$, $p < .001$, partial η^2 = .219, where female faces ($M = .763$, $SE = .079$) were better remembered than male faces (M = .178, $SE = .104$). Additionally a main effect of gender of participant was present ($F(1, 93)$ = 7.496, $p < .01$, partial $\eta^2 = .075$), where women ($M = .670$, $SE = .096$) performed better than men ($M = .202$, $SE = .110$). This replicates evidence suggesting women are better at remembering faces (Shapiro & Penrod, 1986; Wright & Sladden, 2003) No interaction was apparent, $F(1, 93) = .060$.

Items-Level Analysis: One-word mental/emotional state responses that were collected during the encoding phase of the experiment were used as a measure of mentalizing. Two coders blind to the experimental purpose were instructed to code responses as either indicating no mental state, a basic mental state, or a complex mental state. Coders were instructed to code responses that implied no mental state, such as bored, tired, or thinking nothing, as zero. Words that were defined to indicate basic emotions by Ekman (1999) were coded as 1. These included responses such as happy, sad, angry, fearful, surprised, or disgusted, and direct synonyms of these words (e.g., joy for happy, mad for angry, scared for fearful). All other responses were coded as complex mental states and coded as 2. These scores were fairly consistent across the two raters ($r = .806$). Therefore, in order to code one-word attributions to these faces, the two raters' scores were averaged and combined into a single score for each face in each condition. This measure yielded a marginally significant effect ($t(190) = 1.902$, $p < .06$), such that responses to ingroup faces ($M = 1.263$, $SE = .018$) were coded as conveying more complex mental states than were outgroup faces ($M = 1.216$, $SE = .016$). However, there were no group differences in recognition performance ($t(190) = - .507$).To demonstrate a mediation, first, an effect must exist between the independent variable (group status) and the dependent variable (memory) (Judd, Kenny, & McClelland, 2001). Since this study failed to replicate this effect, this study could not test if mentalizing mediated own-group differences in memory.

Discussion

Study 2 replicated the relationship between nonverbal sensitivity and face memory by showing a correlation between the DANVA and recognition memory for faces. This study, however, again failed to show a significant correlation between the Eyes Test and face memory, though this time there was a near trend toward a positive relationship and an effect size comparison revealed no significant differences between the DANVA and the Eyes test in predicting face memory. That said, based on an assumed power of .8 and the effect size of the relationship between the Eyes Test and memory found in this study, it would take over 150 subjects to yield a significant relationship.

Experiment 2 revealed a trend toward an effect for coded complexity scores, as participants attributed more complex mental states to faces that were described as Penn State students than Michigan students. However, Experiment 2 failed to replicate the findings of Bernstein et al. (2007), revealing no ingroup bias in face memory. This was likely not an effect of a failed manipulation since the ingroup faces did evoke greater mentalizing.[1] This finding is also problematic for the hypothesis that mentalizing is associated with face memory, since there was no correlation between the two constructs nor was there a group effect in face memory.

It is possible that mentalizing is not associated with memory when participants are actively asked to mentalize about a face, but there may still be an association between mentalizing and memory when participants are not actively engaged in a mentalizing task. Neuroimaging research shows that the neural correlates underlying emotional processing are distinct based on the cognitive demands of a task (Phan, Wager, Taylor, & Liberzon, 2002). For instance, tasks requiring attentional effort, such as labeling emotions, cause a reduction in the degree to which emotional faces activate the amygdala when compared to merely passively viewing the faces, (Costafreda, Brammer, David, & Fu, 2008). This suggests that it is possible the relationship between mentalizing and emotion may be distinct if a participant is asked to label one's mental state versus a task where a person is not told to label a mental state. However, such an assertion has never been tested. Additionally, Experiment 2 attempted to correlate an implicit memory process with an explicit mental state judgment. It is also possible in this case that the implicit memory processes operate separately from explicit mental processes and it is only the latter that is affected by an explicit mental state task. This is further explored below.

Experiment 3

Experiment 2 failed to replicate Bernstein et al. (2007) using a cross-group face memory paradigm. The only major difference between this experiment and the original Bernstein et al. experiments was that Experiment 2 employed an incidental memory task while Bernstein et al. used an intentional memory task such that participants were told to remember faces for a subsequent recall task. The memory task in Experiment 2 was incidental so the participants could engage in the explicit task of mental state decoding instead. Given that research examining the own-race bias has found cross-race memory effects using both intentional and incidental memory tasks (Meissner & Brigham, 2001; Sporer, 2001), I did not anticipate that this would be an issue. Even when using perceptually ambiguous groups, such as sexual orientation, participants show an ingroup advantage using incidental memory tasks (Rule et al., 2007). If the own-race memory bias and other arbitrarily assigned group memory biases are caused by the same social cognitive mechanism as current models contend (see Sacco & Hugenberg, 2008), then the own-group bias for arbitrarily assigned groups should presumably be present for both incidental and intentional memory tasks. However, this assumption has not been directly examined using arbitrarily assigned groups to my knowledge, so it is unclear whether this may be a contributing factor to why Experiment 2 failed to replicate this previously reported effect.

Intentional and incidental memory tasks are qualitatively different cognitive tasks, which may be a reason why Experiment 2 did not replicate the ingroup bias found previously by Bernstein et al., (2007) using a virtually identical task in all other respects. In intentional memory tasks, participants are able to use encoding strategies and selectively allocate attention for memory in order to remember the items (Fletcher, Frith, & Rugg, 1997, Naveh-Benjamin et al., *in press*). However, incidental memory tests do not allow participants to use explicit elaboration strategies since they have no explicit motivation to encode the faces. Some neuroimaging studies suggest that cognitive resources are reallocated from brain regions implicated in the default network to active tasks in intentional memory paradigms versus incidental memory paradigms (Daselaar, Prince, & Cabeza, 2004; Turk-Browne, Yi, & Chun, 2006).

Intentional memory tasks share activation in some of the same brain networks involved in memory as incidental memory tasks (e.g. Stark & Okado, 2003). However, hippocampal activation in memory tests is specific to intentional memory paradigms (Strange, Hulemann, Duggins, Heinze, & Dolan, 2004). Instructions to remember stimuli cause increases in brain activation in the ventrolateral prefrontal cortex, which supports the idea that there is an additional "top-down" cognitive control component to intentional memory (Dove, Brett, Cusack, & Owen, 2006). These findings provide considerable support for the assertion that intentional memory tasks are functionally distinct from incidental memory by demonstrating different neural correlates for the two processes. Therefore, it is possible that differences between intentional and incidental memory may moderate the effects of social categorizing on memory. The aim of Experiment 3 was to examine this possibility.

Participants: Thirty-two women and 24 men completed this experiment for a total of 56 participants.

Stimuli: The same faces were used as before, only that the manipulation was changed slightly to match the manipulation used by Bernstein et al. (2007). For the ingroup faces, faces were displayed with a dark blue background with "Pennsylvania State University" in large white letters below the face. The outgroup faces were displayed with a yellow background with "University of Michigan" in dark blue letters.

Design and Procedure: The same design and procedure was followed as in Study 1 with an encoding block, a brief four minute distracter task, followed by a retrieval task. Participants were instructed that these faces represented members of either Penn State University or the University of Michigan and that they would have to remember these faces for an upcoming recall task where they would be mixed in with other faces of people from the two schools. Faces were divided into two groups, one representing the target faces while the second group represented the distracter faces. Faces were counterbalanced as to which group they represented, so that half of the participants saw a face as an ingroup face while the other half saw the face as an outgroup face. Faces were displayed for three seconds during the encoding phase in random order.

After the distracter task, participants completed a recognition task like in Experiment 2, where they used the keyboard to indicate if they had previously seen the face before or if the face was a new face. Since Experiments 1 and 2 both showed a relationship between memory and nonverbal sensitivity, no measures of nonverbal sensitivity were administered in this experiment.

Results

Subject Analysis: Participants had a mean overall accuracy of 62.7% on the task, which is consistent with the overall accuracy found in Study 2. Signal detection scores was used to compute recognition performance (overall d' = .706, SD = .396). The two counterbalanced groups of participants did not differ in overall accuracy ($t(54)$ = .394), indicating that there were no effects of memory based on which participants saw which faces in each of the randomly assigned groups.

Using d' scores, a significant difference was found between ingroup and outgroup performance ($t(55)$ = 3.072, p < .005, d = .828), where ingroup performance (M = .839, SE =

.071) was significantly higher than outgroup performance (*M* = .611, *SE* = .062). This indicates that the arbitrary group manipulation significantly influenced memory in the predicted direction.

The effects of gender were by computing a 2(gender of face) X 2(gender of participant) mixed-design ANOVA using again using *d′* as the DV. This revealed main effects of gender of face ($F(1, 54) = 5.282$, $p < .05$), where female faces (*M* = .791, *SE* = .068) were better remembered than male faces (*M* = .633, *SE* = .062). Additionally a trend toward a main effect of gender of participant was present ($F(1, 54) = 2.973$, $p < .09$), where women (*M* = .807, *SE* = .072) performed better than men (*M* = .617, *SE* = .083). No interaction was apparent, $F(1, 54) = .130$.

Discussion

Experiment 3 replicated the ingroup bias that was found by Bernstein et al. (2007) using an intentional memory paradigm. Interestingly, this provides evidence that the own-group bias for arbitrarily assigned groups may have distinct processes as own-race bias or other ingroup biases. Memory advantages for own-race faces extend to incidental memory, especially in real-world contexts where people do not know that they will be tested on what they have seen (such as a potential witness to a crime; Meissner & Brigham, 2001). They also extend to perceptually ambiguous groups, such as sexual orientation. Sexual orientation is a socially salient category but it represents a perceptually ambiguous group, in that there are no obvious or consistent markers, though people do show higher than chance accuracy at recognizing sexual orientation in others (Rule, Ambady, Adams, & Macrae, 2008). Even though sexual orientation is not perceptually distinct, facial cues indicating sexual orientation still affect how male faces are processed by male heterosexual and homosexual participants.

If the same social categorization processes were sufficient to explain the own-race bias and own-group biases more generally, then the own-group bias should be present in all conditions where own-race biases are found, including incidental memory. However, the absence of this finding suggests that an additional process is at work. Differences in social categorization between race and arbitrarily assigned groups may explain why the arbitrary group assignment paradigm used in Experiment 3 and by Bernstein et al. (2007) only showed a cross-group effect for intentional memory, but not for incidental encoding (Experiment 2).

From well-practiced and socially salient categories such as race, age, and gender, to arbitrarily assigned groups, perceivers categorize people based on explicit labels that describe others (e.g., Fiske & Neuberg, 1990). In addition, group-based processing for many socially relevant groups that are perceptually distinct is moderated by bottom-up features of the stimulus itself (e.g., MacLin & Malpass, 2001; Rule et al., 2008). In this case, systematic variation in features that are part of the stimulus (e.g., hairstyle) can affect how a face is perceived. Supporting this distinction, Ito and Urland (2003) found that faces of different races and genders resulted in larger magnitudes of the P300 component in an event-related potential study, which they attributed to explicit categorization effects (see also Ito, Thompson, & Cacioppo, 2004). However, they also found larger responses to other-race faces in the N100 component, which has traditionally been considered a marker of attention, and is thought to occur well before active categorization processes. This strongly supports the assertion that implicit and explicit categorization processes each influence processing of other-race faces, yet it is difficult to envisage a system where arbitrarily assigned categories with only semantic markers of category membership can affect implicit categorization only 100 milliseconds after stimulus perception.

Thus, it is very possible that implicit and explicit social categorization processes each contribute to the own-race bias, yet only explicit processes contribute to own-group biases in arbitrarily assigned groups. Ito and Urland (2003) assert that differential responses in early processing, such as in the N100 component, reflect attentional differences in processing other-race faces. It is possible that this reflects early initial attention to implicitly processing if outgroup faces are threatening or not (see Ackerman et al., 2006), but after this, attention is allocated toward ingroups, who are individuated (see also Dickter & Bartholow, 2007). Thus, this additional attention allocated toward ingroups, even before explicit categorization ostensibly occurs, might cause differences in processing ingroup versus outgroup faces in intentional and incidental memory paradigms. However, it is possible that explicit categorization processes between ingroups and outgroups only affect the strategies used in encoding ingroup or outgroup faces of arbitrarily assigned groups, while not affecting initial attentional differences to these groups. However, to my knowledge, no study has examined the timecourse of categorizing minimal group members, so this represents another question for future research.

Experiment 4

Experiment 3 replicated the ingroup bias in face memory previously reported by Bernstein et al (2007) using the same paradigm. However, this study did not examine the primary hypothesis of this thesis, that being that mentalizing not only affects how faces are remembered but mediates own-group biases in face memory. Although I was unable to replicate the intergroup memory effect using the incidental memory paradigm employed in Experiment 2, I was able to replicate this effect in Experiment 3 using an intentional memory paradigm. Therefore, the hypothesis that categorization leads to less mentalization about a face, and thus less subsequent memory, remains plausible. Experiment 4 sought to replicate Experiment 3, while also testing if mentalizing mediates the relationship between group membership and memory.

Experiment 2 used a methodology where participants mentalized about a face and then subsequently were tested on memory of those faces. That is the most direct way in order to assess the relationship between mentalization and face memory. However, it requires an incidental memory task and for reasons described above, the effect of group membership on memory may be moderated by if a memory test is intentional or incidental. Since Experiment 1 showed that a scale of mentalization predicted memory in the same way as did the one-word attribution task, the present experiment tested the relationship between mentalizing and memory by using two separate tasks. First, participants completed an intentional memory task in the same way as described in Experiment 3. Second, these same participants rated the same faces on the 7-point mentalization scale described in Experiment 1. This scale was used to see if participant differences in mentalization for ingroup versus outgroup faces mediated memory.

Method

Participants: Fifty-four participants (34 female, 20 male) completed the study from the undergraduate psychology subject pool.

Stimuli: The same stimuli with the same group manipulation were used as in Study 3.

Design and Procedure: The experiment consisted of two parts. First, participants completed the same explicit memory task as in Experiment 3 where half the faces were shown as ingroup (Penn State) faces and the other half as outgroup (Michigan) faces. Following the memory task, participants completed the mentalizing scale. For the mentalizing scale, participants were instructed that they would see the same faces as before and would rate them on a seven point scale on how complex of a mental state they had. Participants were told that if the face appeared to have no mental state to rate the face low on the scale, if the face appeared to have a basic emotional state, such as anger, happiness, or fear, to rate the face in the middle of the scale, and if the face had a complex mental state, to rate the face highly in the scale. These were the same instructions used in Phase 2 of Experiment 1.

Faces were counterbalanced so that one group of the subjects saw half the faces in the ingroup context while the other group of subjects saw those same faces in the outgroup context. Faces were always shown in the same group context in both the encoding and recall phases, as well as in the mentalizing scale.

Results

The hypothesis of interest tested in this study was that the ingroup advantage in face memory is mediated by differences in the extent to which people mentalize about ingroup versus outgroup faces. To demonstrate a mediation, first, an effect must exist between the independent variable (group status) and both the dependent variable (face memory) and mediator variable (mentalizing) (Judd, et , 2001).

To test the group advantage in memory, signal detection scores (*d′*) were computed separately for ingroup and outgroup faces. This revealed an ingroup advantage, $t(53) = 2.976$, $p < .005$, $d = .818$, where ingroup faces ($M = .928$, $SE = .057$) were remembered better than outgroup faces ($M = .729$, $SE = .056$). Likewise, ingroup faces ($M = 3.29$ out of 7, $SE = .105$) received a higher score on the mentalizing scale than outgroup faces ($M = 3.21$, $SE = .101$); $t(53) = 2.074$, $p < .05$, $d = .570$.

Given that the independent variable had an effect on both the mediator as well as the dependent variable, the next step to determining within-subject mediation involves calculating difference scores for the mediator and dependent variable as a function of the independent variable (Judd, Kenny, & McClelland, 2001). Since the independent variable of interest (i.e.,. perceived group membership, ingroup versus outgroup) yielded significant effects for both the dependent variable and mediator, difference scores were computed for each by subtracting each participant’s outgroup measure from their ingroup measure. To determine mediation, the difference scores for the memory test were then regressed onto the difference scores of the mentalizing scale. If this difference is significant, then within-subjects mediation has been indicated. This analysis revealed a marginal effect in the direction of a mediation, ($B = .197$, $p = .079$ one-tailed).

Discussion

This study replicated Experiment 3 as well as Bernstein et al. (2007) in finding an ingroup advantage in face memory in an intentional memory task. This provides additional support for the hypothesis suggested above that the own-group bias may only be present for intentional memory tasks. Additionally, this study replicated evidence that an own-group bias exists in mentalizing, as found in Experiment 2, suggesting that people read more complex mental states into faces that are presented as ingroup members versus outgroup members. Since these faces were counterbalanced for each condition, this cannot be an effect of stimulus, as each face was seen an equal amount of times as an ingroup versus an outgroup face.

Experiment 4 also revealed initial evidence in the form of a trend for the mediation hypothesis that was proposed. This suggests that on a subject-level, mentalization might mediate the relationship between group membership and face memory, but such a weak effect warrants replication before any firm conclusions can be drawn. In this case, the next step may be to replicate the condition in Experiment 2 where participants engage in an explicit mental state decoding task about faces while fully knowing that those faces are to be remembered for an upcoming memory test. In this case, both the mental state decoding condition and the memory test will be explicit tests, addressing the potential weakness in Experiment 2 where mental state decoding was explicit while the memory task was implicit.

Another possible weakness of this mediation design is that the mediator variable was measured after the dependent variable, which violates the normal temporal order that is necessary for mediation. However, in this case, the mediating variable is hypothesized to occur before the dependent variable, yet it is being measured afterward as a proxy measurement. In this case, the mediator variable can be measured after the dependent variable (see Goff, Steele, & Davies, 2008; Keller & Dauenheimer, 2003). In this case, the mentalizing scale was done as an exploratory analysis since I hypothesize it is part of the process in viewing a face. Because of this, and because the findings were only marginally significant, the mediation analysis is reported only as a pilot analysis and warrants further replication.

Chapter 3: General Discussion

This thesis examined the relationship between mentalizing, or the process of reading mental states from faces, and subsequent face memory. Previous models of face perception proposed that these are cognitively distinct processes with separate neural underpinnings (Bruce & Young, 1986; Haxby et al., 2000; but see Adams et al., in press; Calder & Young, 2005). However social psychological literature on group dynamics demonstrates robust ingroup biases as people associate more complex mental states and remember better those people of their own group. These parallel findings suggest that mentalizing and memory may be related processes. Thus, this thesis examined the relationship between mentalizing and memory as it pertains to three factors, stimulus factors, perceiver factors, and relational factors.

The relationship between mentalizing and face memory on the level of perceiver factors was examined by studying the relationship between individual differences in face memory and individual differences in mentalizing, as measured by nonverbal sensitivity. These revealed a relationship in both Studies 1 and 2 indicating that people who were better at discriminating low-intensity basic emotions, as measured by the DANVA, were better at remembering faces. This correlation suggests a relationship between the two processes, providing to my knowledge the first evidence that nonverbal sensitivity and face memory are related.

This thesis also found support for the relationship between memory and mentalizing at a stimulus level. I hypothesized that some faces would evoke more mentalization than other faces. That is, there is something about some faces that makes a perceiver desire to read more into the mental states of certain faces more than others, even though all of those faces ostensibly have no differences in the expressive mental states that they are conveying. This thesis found robust support for this, as people showed high agreement as to the degree to which they attributed basic versus complex mental states to faces. Additionally, those faces that one group of raters attributed more complex mental states to were better remembered by another group of participants while using an intentional memory test. This suggests that there exists a quality intrinsically present within a face that evokes more mentalization. This quantity may be related to other known stimulus factors that affect person perception, such as attractiveness, babyfaceness, dominance, or affiliation, but it represents a separate mechanism. The relationship between this factor and memory may be driven by the fact that certain faces cause people to read more into them to determine their mental state and thus are deeper processed and are better encoded. Future studies are needed to ascertain the mechanism by which this process happens and what features are associated with this factor that causes participants to rate faces as having more complex mental states.

This thesis also investigated the relationship between mentalizing and face memory as it involved the own-group bias. Literatures suggested that the own-group bias is due to differential categorization of ingroup and outgroup faces, where outgroup faces are categorized and thus not processed as deeply, while ingroup faces are individuated to a greater degree and are more deeply processed. Since there also exists a relationship between mentalization and group membership, I hypothesized that mentalization may mediate the own-group bias as the difference in reading mental states into faces predicted different memory of those faces. Experiment 4 hinted that such a relationship may exist. However, discrepancies in the own-group bias for incidental versus intentional encoding suggest that the social categorization mechanism theorized to be behind the own-group bias may be distinct than that mechanism behind perceptually distinct groups.

Incidental versus Intentional Memory and Own-Group Biases

An unexpected, but interesting finding in this thesis was a failure to replicate the own-group bias in Experiment 2, using an incidental memory task. Bernstein et al. (2007) and Experiments 3 and 4 replicated this using an intentional memory task, showing that such an effect was likely an effect of the task instructions, which was the only significant difference between the two paradigms. This suggests that the own-group bias for faces arbitrarily assigned to an ingroup versus an outgroup university might represent a different process than the own-race bias.

One potential explanation for this finding may be that differences in categorization exist between perceptually distinguishable groups and perceptually indistinguishable groups such as university affiliation. Categorization based on group membership is often considered as an explicit process modulated by semantic, category-based information. However, categorization also is an implicit process that occurs without conscious control (Ito & Cacioppo, 2000; Wittenbrink, Judd, & Park, 1997). Implicit categorization, as discussed above, is based on perceptual category variation within the stimuli themselves, occurring before semantic category activation can be active (Ito and Urland, 2003) It is possible that categorization is a feature-based process motivated by differences in the cues that are learned. For instance, cues such as identity (Bruce & Young, 1998) and gender (Cloutier, Mason, & Macrae, 2005) can be extracted efficiently and under conditions of high cognitive demand. It is also the case that irrelevant categorical information is gleaned from processing as well, as long as participants are motivated to process some categorical information from a face (Macrae, Quinn, Mason, & Quadflieg, 2005).

If implicit categorization moderates face perception based on attentional differences, then implicit categorization should affect memory in incidental and intentional memory paradigms.

This would be true as long as participants are given a face-relevant processing task since Macrae and colleagues (2005) found that was all that was sufficient to give rise to implicit categorization processes. Since there are no group-based perceptual differences between group members in the arbitrarily assigned group task, then it is unlikely that this group manipulation would affect implicit feature-based categorizations. On the other hand, an arbitrarily assigned group manipulation may only impact the explicit processes that are especially involved in intentional memory, such as encoding strategies, since those are more explicit and consciously controlled. Therefore, so while the own-race bias may be driven by differences in implicit and explicit social categorization of stimuli, the own-group bias found by Bernstein et al., (2007) and in my Experiments 3 and 4 may only be driven in explicit categorization differences.

Stereotypes may play a role in how this implicit categorization is learned. The features that give rise to categories have to be learned at some point in development and it is possible that the stereotypes relevant to categories interact with how we learn to categorize. Associations can be learned between faces that vary on a facial dimension and randomly associated personality traits that covary with that dimension in a mere laboratory setting (Hill, Lewicki, Czyzewska, & Schuller, 1990). Thus, years of development based on categories with associations between group stereotypes and facial features (e.g. features common to Black faces and Black stereotypes) could help give rise to a more automatic implicit process of categorization. Race and sexual orientation are two groups associated with robust stereotypes and these groups show own-group biases in memory for incidental and intentional tasks. However, groups such as university affiliation are arguably not associated with long-term stereotypes and are very likely not associated with meaningful covariance in facial dimensions. Thus, they are only affected by explicit categorization processes. This represents another line of research that needs to be examined in the future.

Future Directions

Manipulations of Mentalizing and Subsequent Memory: A stronger test of the relationship between mentalizing and memory would be a manipulation of the degree to which a participant mentalized about a face and the degree to which that face was remembered. There are two ways that I plan to test this hypothesis. First, I plan to test a direct manipulation of mentalizing by using a paradigm similar to Mitchell, Macrae, & Banaji (2006). In this, faces would be displayed in two conditions. First, participants would see faces in a condition where they would only make a perceptual-level judgment about the face, such as if the face was symmetrical. In the mentalizing condition, they would be asked to rate a face based on what it is thinking or feeling, using a similar wording as used in the mental state complexity scale above. In this case, I hypothesize that the second condition would force participants to mentalize about a face more and thus show better memory.

The second way in order to test this hypothesis is to use a more constrained mentalizing paradigm. There are situations, for instance, where we are motivated more to ascertain if a person if feeling a certain way rather than to read what way that a person is feeling. For instance, if we expect a person to possibly be threatening, it would be more adaptive to see if a person is angry rather than to read the mental state with no expectation. Therefore, the way in which I plan to test if mentalizing affects memory in a constrained mentalizing task is testing if reading if a face has basic versus complex mental state words affect how well that face is subsequently remembered. In this case, participants would be asked to indicate the degree to which a neutral face matches words that are either basic emotional words or complex mental states (for a similar paradigm, see Wenger & Ingvalson, 2002). Afterward, I expect in a surprise memory test that the faces that were matched with complex mental states would be better remembered than those matched with basic emotions. These two experiments would indicate that mentalizing affects memory in two different manipulations.

Mentalizing as a Mediator of the Own-Race Bias: The hypothesis tested in Experiment 4 that mentalizing mediates the own-group bias received very limited support. However, this hypothesis was drawn from the literature that suggested parallel own-race biases in memory and in attributing complex mental states to other races. Theorists argued that each of these biases may be group-related and generalize to any ingroup versus outgroup situation (e.g. Hugenberg & Sacco, 2008; Paladino et al., 2002). However, the experiments contained herein argue that own-group biases may be different than the ubiquitous own-race bias. This leads to the conclusion that while the mediation model for the own-group bias only received marginal support, this mediation model may still be supported in the own-race bias.

An experiment to analyze this conclusion would proceed in the manner of Experiment 2, where mentalizing is directly analyzed as a mediator of face memory. In this case, I would expect an own-race bias in the incidental memory task. Additionally, I would also expect that participants would attribute more complex mental states to same-race faces. This finding is suggested by the literature and by research in our laboratory, but no published studies to my knowledge have shown that more complex mental states are attributed to neutral faces of the same race as opposed to other-race faces. In this case, if mentalizing mediates the own-race bias, then this would suggest that level of mentalizing may be part of the explanation of why same-race faces are remembered better.

Differences Between the Own-Race Bias and Own-Group Biases: Recent evidence has made a compelling argument that the own-race bias is not a perceptual phenomenon, but rather is based on the social categorization of the stimuli (Hugenberg & Sacco, 2008; Levin, 1996; 2000). However, this research suggests that these biases may represent different processes. As noted above, differences in implicit versus explicit categorization processes could explain discrepant findings between the own-race bias and own-group biases in intentional versus incidental memory tasks. First, I expect that implicit categorization is more dramatically affected in groups that are perceptually distinct. This would occur in addition to biases due to social categorization. Additionally, it is also possible that learned stereotypes may affect how this degree of implicit categorization is learned.

Mechanisms exist to test both of these hypotheses. If the first one is true, that implicit and explicit categorization processes affect the own-race bias, it should be the case that in an intentional memory test, measures of implicit and explicit categorization processes each should partially mediate the own-race bias. However, in an incidental memory test, only implicit categorization should affect the own-race bias.

The second hypothesis is that learned stereotypical associations associated with groups such as race and sexual orientation give rise to incidental memory effects for these groups. This is not well-supported by the literature, as Meissner and Brigham (2001) found no effect between racial attitudes and the own-race bias in a meta-analysis. However, most of these studies used intentional memory tests, which may be motivated by social categorization independent of racial attitudes. Additionally, it is possible that this lack of an effect could be due to imperfect measures of racial attitudes, as suggested by Sporer (2001). Indeed, implicit racial bias as measured by the implicit association test (IAT) correlates with the memory of outgroup faces

(Walker & Hewstone, 2008). Therefore, if this hypothesis is true, learned stereotypical associations based on implicit racial attitudes would mediate the own-race bias in incidental memory tests. Therefore, to examine this, an experiment would test to see if IAT scores mediate the own-race bias in intentional and incidental memory tests. If learned stereotypes cause the effect found herein between incidental and intentional memory, this study should find that the IAT scores should mediate the own-race bias in incidental memory much more than intentional memory.

Neural Underpinnings of the Relationship between Mentalizing and Memory: Another future direction that this research suggests is an examination into the neural underpinnings of the relationship between mentalizing and face memory. More mentalizing leads to greater facial memory. However, the neural underpinnings of such a relationship could help uncover the cognitive mechanism by which greater mentalizing causes greater memory. Specifically, a study that compared the activation between faces that evoke high levels of mentalization versus those that evoke low levels while doing a memory test would examine how this stimulus-level variation affects memory. Though there are many different brain regions that could reflect differences between high-mentalizing and low-mentalizing faces, one of note would be the superior temporal sulcus (STS). This region is consistently involved in processing dynamic aspects of the face, such as eye gaze (Hoffman & Haxby, 2000), facial movements (Puce, Allison, Bentin, Gore, & McCarthy, 1998), and even mental state decoding (Adams et al., in press). If neutral faces that vary in mentalizing differentially activate the STS in a memory task, and if that STS activation predicts greater memory for high versus low mentalizing faces, that would provide compelling support that mentalizing is tied into the way faces are remembered.

Conclusion

The research presented in this thesis represents a novel way in considering the relationship between mentalizing and face memory. Within the four studies presented here, converging support is found that a relationship exists. This is strongest when studied in the context of stimulus variation, as those faces which evoke more mentalization are better remembered. Additionally, face memory correlated with scores on the DANVA, a nonverbal sensitivity measure. The predicted mediation model, that mentalizing would mediate the own-group bias, received only slight support. Even though this failed to find strong support for a relationship, the data contained herein provides some of the initial groundwork to explore how reading mental states into faces affects how we remember and process those faces.

The other interesting finding was the failure to replicate the own-group bias using an implicit memory measure. This provides evidence that the own-race bias may be more complex than what can be explained by social categorization measures and that the own-group bias and the own-race bias may not be parallel effects with the same underlying mechanism. Since this was an underlying assumption to the mediation model tested in this thesis, this finding does not preclude that mentalizing is unrelated to own-group biases. Indeed, as mentioned above, nothing found within this thesis suggests that mentalizing does not mediate the own-race bias. Again, this is a question for future research.

This thesis represents an initial step in delineating the relationship between mentalizing and memory. Such a relationship offers unique insight into how the brain processes faces and perhaps why faces are processed in the way that they are. This thesis also suggests that the social meaning of stimuli must be taken into account in considering how such stimuli are processed. Faces have often been said to be “special” as stimuli or in the way they are processed. Though this has been debated, it is much less debatable about how special the face is in the amount of social meaning that it conveys. Reading people's mental states is critically important in social interaction. Therefore, it is not surprising that mentalizing interacts with how people remember faces, and this thesis represents a first step in determining the interaction of these two previously considered disparate processes.

Bibliography

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V. et al. (2006). They all look the same to me (unless they're angry): From out-group homogeneity to out-group heterogeneity. *Psychological Science*, *17,* 836-840.

Adams, R. B. Jr., Pauker, K. & Ambady, N. (2008). Eye gaze influence race-related memory for faces. *Manuscript in Preparation*.

Adams, R.B., Jr., Rule, N, Franklin, R.G., Jr., Wang, E., Stevenson, M.T., Yoshikawa, S., Nomura, M., Sato, W., Kveraga, K., & Ambady, N. (in press). Cross-cultural reading the mind in the eyes. An fMRI investigation. *Journal of Cognitive Neuroscience*.

Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience, 14*, 1264-1274.

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C., (2001). Beatuiful faces have variable reward value: fMRI and behavioral evidence. *Neuron, 32*, 537-551.

Allison, T., Puce, A. & McCarthy, (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4*, 267-278.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268-277.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6*, 269-278.

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., & Williams, S. C. R. (1999). Social intelligence in the normal and autistic brain: An fMRI study. *European Journal of Neuroscience, 11,* 1891-1898.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. (2001). The "Reading of the Mind in the Eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high functioning autism. *Journal of Child Psychology and Psychiatry, 42*, 241-251.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*, 706-712.

Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyish face. *Journal of Personality and Social Psychology, 48*, 312-323.

Blair, I. V. & Judd, C. M. (in press). Afrocentric facial features and stereotyping. Chapter forthcoming in R. B. Adams Jr., N. Ambady, K. Nakayama, & S. Shimojo (Eds.) *The Science of Social Vision*.

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology, 87,* 763-778.

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology, 83*, 5-25.

Brigham, J. C., & Ready, D. J. (1985). Own-race bias in lineup construction. *Law & Human Behavior, 9,* 415–424.

Calder, A. J. & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*, 641-651.

Castelli, F., Happè, F., Frith, U. & Frith, C. (2005). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. In J. T. Cacioppo & G. G. Berntson (Eds.) *Social neuroscience: Key readings: Key reading in social Psychology* (pp. 155-169). New York: Psychology Press.

Cloutier, J., Mason, M. F., & Macrae, C. N. (2005). The perceptual determinants of person construal: Reopening the social-cognitive toolbox. *Journal of Personality and Social Psychology*, *88*, 885.

Chance, J., Goldstein, A., & Mcbride, L., (1975). Differential experience and recognition memory for faces. *Journal of Social Psychology*, *97*, 243-253.

Corden, B., Critchley, H. D., Skuse, D. & Dolan, R. J. (2006). Fear recognition ability predicts differences in social cognitive and neural functioning in men. *Journal of Cognitive Neuroscience*, *18*, 889-897.

Corneille, O., Hugenberg, K., & Potter, T. (2007). Applying the attractor field model to social cognition: Perceptual discrimination is facilitated but memory is impaired for faces displaying evaluatively congruent expressions. *Journal of Personality and Social Psychology*, *93*, 335-352.

Costafreda, S. G., Brammer, M. J., David, A. S., & Fu, C. H. Y. (2008). Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies. *Brain Research Review*, *58*, 57-70.

Craik, F. I. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.

Cuddy, A. J. C, Rock, M. S., & Norton, M. I. (2007). Aid in the aftermath of Hurricane Katrina: Inferences of secondary emotions and intergroup helping. *Group Processes and Intergroup Relations, 10*, 107-118.

Daselaar, S. M., Prince, S. E. & Cabeza, R. (2004). When less means more: Deactivations during encoding that predict subsequent memory. *NeuroImage, 23*, 921-927.

Depaulo, B. M. & Friedman, H. S. (1998). Nonverbal communication. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.). *The handbook of social psychology, 4th ed*, (pp. 3-40). New York: McGraw-Hill.

Eberhardt, J. L., (2006). Imaging race. *American Psychologist, 60*, 181-190.

Ekman, P., (1999). Basic emotions. In T. Dalgleish and M. Power (Eds.). *Handbook of Cognition and Emotion.* Sussex, U.K.: John Wiley & Sons, Ltd., 1999.

Ekman, P., & Friesen, W. (1975). Unmasking the Face. A Guide to Recognizing Emotions from Facial Expressions. Consulting Psychologists Press, Palo Alto, CA.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is special about face perception? *Psychological Review, 105*, 482-498.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist, 48*, 621-628.

Fiske, S. T., & Neuberg, S. L. (1990). A continum of impression formation, from category ? based to individuating processes: influences of information and motivation on attention and interpretation. *Advances in experimental social psychology*, *23*, 1-74.

Fletcher, P. C., Frith, C. D. & Rugg, M. D. (1997). The functional neuroanatomy of episodic memory. *Trends in Neuroscience, 20*, 213-218.

Franklin, R. G., Jr. & Adams, R. B., Jr. (under review-a). A dual-process account of female facial attractiveness preferences: Sexual and nonsexual routes.

Franklin, R. G., Jr. & Adams, R. B. Jr. (under review-b). The Two Halves of Beauty: The Laterality of a Dual-Process of Facial Attractiveness.

Franklin, R. G., Jr. & Adams, R. B., Jr. (under review-c). The interactive nature of eye gaze and emotion in face memory.

Frith, U. & Frith, C. (2001). The biological basis of social interaction. *Current Directions in Psychological Science, 10*, 151-155.

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.

Goff, P. A., Steele, C. M., & Davies, P. G. (2008). The space between us: stereotype threat and distance in interracial contexts. *Journal of Personality and Social Psychology, 94*, 91-107.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 85*, 845-857.

Hall, J. A., Carter, J. D., & Horgan, T. G. (2001). Status roles and recall of nonverbal cues. *Journal of Nonverbal Behavior, 25*, 79-100.

Hall, J. A., Coats, E. J., & Smith LeBeau, L. (2005). Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin, 131*, 898-924.

Henley, N. M. (1977). *Body politics: Power, sex, and nonverbal communication.* Englewood Cliffs, NJ: Prentice-Hall.

Hess, U., Blairy, S., & Kleck, R. E., (2000). The influence of facial emotion displays, gender, and ethnicity on judgments of dominance and affiliation. *Journal of Nonverbal Behavior, 24*, 265-283.

Hill, T., Lewicki, P., Czyzewska, M., & Schuller, G. (1990). The role of learned inferential encoding rules in the perception of faces: Effects of nonconscious self-perpetuation of a bias. *Journal of Experimental Social Psychology, 26*, 350-371.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience, 3*, 80-84.

Hugenberg, K. & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14*, 640-643.

Hugenberg, K. Miller, J., & Claypool, H. M., (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology, 43*, 334-340.

Hugenberg, K. & Sacco, D. F., (2008). Social categorization and stereotyping: How social categorization biases person perception and face memory. *Social and Personality Psychology Compass, 2*, 1052-1072.

Ito, T. A., & Cacioppo, J. T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology, 36*(6), 660-676.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616-626.

Ito, T. A., Thompson, E., & Cacioppo, J. T. (2004). Tracking the Timecourse of Social Perception: The Effects of Racial Cues on Event-Related Brain Potentials. *Personality and Social Psychological Bulletin, 30*, 1267-1280.

Jaquet, E., Rhodes, G., & Hayward, W. G. (2007). Race-contingent aftereffects suggests distinct perceptual norms for different race faces. *Visual Cognition, 16*, 734-753.

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation in moderation in within-subjects designs. *Psychological Methods, 6*, 115-134.

Judd, C. M., & Park, B. (1992). Outgroup homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology, 54*, 778-788.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302-4311.

Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin, 29*, 371-381.

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior, 20*, 165-182.

Lee, K., Byatt, G. & Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: Testing the face-space framework. *Psychological Science, 11*, 379-385.

Leopold, D. A., O'Toole, A. J., Vetter, T. & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience, 4*, 89-94.

Leopold, D. A., Rhodes, G., Müller, K.-M., & Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society of London, Series B, 272*, 897-904.

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1364–1382.

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General, 129*, 559–574.

Lewin, C. & Herlitz, A. (2002). Sex differences in face recognition – Women's faces make the difference. *Brain and Cognition, 50*, 121-128.

Leyens, J. P., Cortes, B., Demoulin, S., Dovidio, J. F., Fiske, S. T., Gaunt, R. et al. (2003). Emotional prejudice, essentialism, and nationalism: The 2002 Tajfel Lecture. *European Journal of Social Psychology, 33*, 703–717.

Lieberman, M. D. & Rosenthal, R. (2001). Why introverts can't always tell who likes them: Multitasking and nonverbal decoding. *Journal of Personality and Social Psychology, 80*, 294-310.

Lindsay, D. S., Jack, P. C. Jr., & Christian, M. A., (1991). Other-race face perception. *Journal of Applied Psychology, 76*, 587-589.

Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face aftereffects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society of London, Series B, 272*, 2283-2287.

MacLin, O. H. & Malpass, R. S. (2001). Racial categorization of faces: The ambiguious race effect. *Psychology, Public Policy, and Law, 7*, 98-118.

Macrae, C. N., Quinn, K. A., Mason, M. F. & Quadflieg, S. (2005). Understanding others: The face and person construal. *Journal of Personality and Social Psychology, 89*, 686-695.

Malpass, R. S. & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13*, 330-334.

Mason, M. F., Cloutier, J. & Macrae, C. N. (2006). On construing others: Category and stereotype activation from facial cues. *Social Cognition, 24*, 540-562.

Mason, M. F., Hood, B. M., & Macrae, C. N. (2004). Look into my eyes: Gaze direction and person memory. *Memory, 12*, 637-643.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces. *Psychology, Public Policy, and Law, 7,* 3–35.

Naveh-Benjamin, M., Shing, Y. L., Kilb, A., Werkle-Bergner, M., LIndenberger, U., & Li, S. (in press). Adult age differences in memory for name-face associations: The effects of intentional and incidental learning. *Memory*.

Mignault, A. & Chaudhuri, A., (2003). The many faces of a neutral face: Head tilt and perception of dominance and emotion. *Journal of Nonverbal Behavior, 27*, 111-132.

Nowicki, S., Jr. & Carton, J. (1993). The measurement of emotional intensity from facial expressions. *Journal of Social Psychology, 133*, 749-750.

Nowicki, S., Jr. & Duke, M. P. (2001). Nonverbal receptivity: The Diagnostic Analysis of Nonverbal Accuracy (DANVA). In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal Sensitivty: Theory and Measurement* (pp. 183-198). Mahwah, NJ: Lawrence Erlbaum.

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologica, 41*, 147-155.

O'Donnell, C. & Bruce, V. (2001). Familiarisation with faces selectively enhances sensitivity to changes made to the eyes. *Perception, 30*, 755-764.

Ostrom, T. M., & Sedikides, C. (1992). Outgroup homogeneity effects in natural and minimal groups. *Psychological Bulletin, 112*, 536–552.

Paladino, P. M., Leyens, J. P., Rodriguez, R. T., Rodriguez, A. P., Gaunt, R., & Demoulin, S. (2002). Differential association of uniquely and nonuniquely human emotions to the ingroup and the outgroups. *Group Processes and Intergroup Relations*, *5*, 105-117.

Pauker, K., Weisbuch, M., Ambady, N., Sommers, S. R., Adams, R. B. Jr., & Ivcevic, Z. (in press). Not so black and white: Memory for ambiguous group members. *Journal of Personality and Social Psychology*.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16,* 331-348.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience, 18,* 2188-2199.

Rehnman J., & Herlitz, A., (2006). Higher face recognition ability in girls: Magnified by own-sex and own-ethnicity bias. *Memory, 14,* 289-296.

Rhodes, G., Simmons, L. W., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success. *Evolution and Human Behavior, 26*

Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitiivty to nonverbal communication: The PONS test.* Baltimore, MD: Johns Hopkins University Press.

Rosip, J. C., & Hall, J. A. (2004). Knowledge of nonverbal cues, gender, and nonverbal decoding accuracy. *Journal of Nonverbal Behavior, 28,* 267-286.

Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2007). Us and them: Memory advantages in perceptually ambiguous groups. *Psychonomic Bulletin and Review, 14*, 687-692.

Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2008). Accuracy and awereness in the perception and categorization of male sexual orientation. *Journal of Personality and Social Psychology,95*, 1019-1028.

Sczesny, S., & Kühnen, U. (2004). Meta-cognition about biological sex and gender-stereotypic physical appearance: Consequences for the assessment of leadership competence. *Personality and Social Psychology Bulletin, 30*, 13-21.

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*, 139–156.

Shepherd, J. W., & Ellis, H. D. (1973). The effect of attractiveness on recognition memory of faces. *American Journal of Psychology, 86*, 627-633.

Sherif, M. (1956). Experiments in group conflict. *Scientific American, 195*, 54-58.

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the faces: Social context modulates the cross-race effect in face recognition. *Personality and Social Psychology Bulletin, 34*, 260-274.

Sporer, S. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law, 7*, 36-97.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology 1982* (pp. 290-312). Washington DC: American Sociological Association.

Somerville, L. H., Wig, G. S., Whalen, P. J., & Kelley, W. M. (2006). Dissociable medial temporal lobe contributions to social memory. *Journal of Cognitive Neuroscience, 18*, 1253-1265.

Strange, B. A., Hrulemann, R., Duggins, A., Heinze, H. J., & Dolan, R. J. (2005). Dissociating intentional learning from relative novelty responses in the medial temporal lobe. *NeuroImage, 25*, 51-62.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223*, 96-102.

Turk-Browne, N. B., Yi, D. J., & Chun, M. M. (2006). Linking implicit and explicit memory: Common encoding factors and shared representations. *Neuron, 49*, 917-927.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race on face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 43*(A), 161–204.

Valentine, T. (2001). Face-space models of face recognition. In M. J. Wenger & J. T. Townsend (Eds.). *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges*. Hillsdale, NJ: Erlbaum.

Walker, P. M. & Hewstone, M. (2008). The influence of social factors and implicit racial bias on a generalized own-race effect. *Applied Cognitive Psychology, 22*, 441-453.

Wenger, M. J. & Ingvalson, E. M. (2002). A decisional component of holistic encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 872-892.

Willis, J. & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592-598.

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.

Wright, D. B. & Sladden, B. (2003). An own-gender bias and the importance of hair in face recognition. *Acta Psychologica, 114*, 101-114.

Wright, D. B. & Stroud, J. N. (2002). Age differences in lineup identification accuracy: people are better with their own age. *Law and Human Behavior*, *26*, 641-654.

Zebrowitz, L. A. (1997). *Reading Faces*. Boulder, CO: Westview.

Zebrowitz, L. A. (2006). Finally, faces find favor. *Social Cognition*, *24*, 657-701.

Zuckerman, M., Lipets, M. S., Koivumaki, J. H. & Rosenthal, R. (1975). Encoding and decoding nonverbal cues of emotion. *Journal of Personality and Social Psychology*, *32,* 1068-1076.

Notes

[1] Another potential explanation for this finding is the idea that the outgroup used may have been considered threatening. Ackerman et al. (2006) found that making an outgroup threatening eliminated the own-race bias. At the time of Experiment 2, the University of Michigan had won nine straight football games over Penn State, leading some students who were asked about this ingroup manipulation to suggest that Michigan represented a threat. In order to account for this possibility, Experiment 2 was replicated using the University of Minnesota, another football rival, as an outgroup. The University of Minnesota's football team was 1-11 the year prior to running this study, leading to the assertion that this rival would be much less threatening. This revealed no difference between the mean scores for the ingroup condition (M = 1.00, SE = .09) and outgroup condition (M = 1.02, SE = .09; $t(30)$ = -.215). Again, this failed to replicate Bernstein et al. (2007) and further suggests that this effect cannot be explained as merely a lack of power.