THESIS

EFFECTS OF IN-GROUP BIAS ON FACE RECOGNITION USING MINIMAL GROUP PROCEDURES

Submitted by

Maia T. Nguyen

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2014

Master's Committee:

Advisor: Lucy J. Troup

Bruce Draper
Matthew G. Rhodes

ABSTRACT

EFFECTS OF IN-GROUP BIAS ON FACE RECOGNITION USING MINIMAL GROUP PROCEDURES

The current series of experiments examined the effects of social categorization on face recognition. The use of minimal group procedures was expected to enhance recognition for in-group members compared to out-group members. In Experiment 1, participants were assigned to 1 of 3 conditions: name study—participants studied a list of 16 names associated with their in-group [red or green], numerical estimation—participants were randomly divided into 2 groups [red or green] after estimating the number of dots in a series of 10 images, and the control condition. This was followed by a study phase in which participants were presented with a total of 32 female and male Caucasian faces on red or green backgrounds. A final recognition test was given following a filler task. Experiment 2 had two of the previously used conditions, name study and control. Faces were presented on red and green backgrounds during test—with old faces presented on the same background as seen at study. Experiment 3 presented a subset of stimuli used in Experiment 2 with a longer presentation time (10 seconds). Findings suggest only moderate difference in response bias between experimental and control groups overall in Experiments 2 and 3. Moderate differences in hits, false alarms, and d’ were also found in Experiment 3 between experimental conditions. Group membership did not elicit significant effects on measures of accuracy, reaction time, and confidence ratings.

TABLE OF CONTENTS

INTRODUCTION

The study of human facial recognition has a history that is as long as it is broad. Initial work in face recognition focused on determining whether our ability to recognize faces is distinctive or a function of expertise (Diamond & Carey, 1986). The study of how we process faces has continued to spur questions regarding the intricacies of our personal interactions (Bruce & Young, 1986). One intuitive approach to understanding facial recognition is to examine what faces we can more easily recognize. For example, one could compare recognition for individuals within a social group to those outside of the group. This type of approach is not exclusive to facial recognition. A broad range of variables have been measured in response to group membership, such as economic decisions, judgments of trustworthiness, and stereotype applications (Desteno, Dasgupta, Bartlett, & Cajdric, 2004; Tajfel, Billig, Bundy, & Flament, 1971; Williams, 2001). The study of in-group bias in facial recognition has evolved by examining the world through the lens of categorized peoples and our behavior toward them.

Own-Race Bias

One of the most prominent branches of research regarding biases in recognition began with our drive to understand race and how we perceive race in others. Meissner and Brigham (2001) wrote an extensive review of the literature regarding the role of race in face recognition. The review expounds upon the finding that individuals can more quickly and accurately recognize individuals of their own race compared to those of other races, also known as the own-race bias (ORB; (Malpass & Kravitz, 1969). In a typical ORB paradigm, participants study same- and other-race faces and are then asked to complete a recognition test. For example, Meissner, Brigham, and Butz (2005; Experiment 2) had participants study White faces and Black faces. Participants were then given a recognition test to determine if faces were ‘old’—seen during the study phase—or ‘new’—not seen before. The test included equal amounts of White and Black faces as well as an equal number of faces seen during the study phase and not seen during the study phase. In addition to demonstrating evidence for the ORB, Meissner and colleagues suggested that own-race faces are encoded with more qualitative information, which in turn leads to heightened familiarity and recognition. An example of qualitative information in this example would be to process features within a face as part of a whole unit, instead of isolated parts.

Quality of encoding has been posited as a primary explanation for why we recognize own-race faces more effectively. Stahl, Wiese, and Schweinberger (2008) used electrophysiological measures to provide additional support for the uniqueness of the ORB. Specifically, they demonstrated that time-locked scalp electrical activity was delayed and amplified while viewing images of other race faces. The delay in activation, as well as the increased magnitude of scalp electrical activity for out-group faces, suggests that the mechanisms associated with processing own- and other-race faces are unique. This study provided additional evidence that race, a way of categorizing individuals, can change the way we process faces.

Own-Age Bias

While race was and is still a primary method people use to distinguish one another, it is not the sole differentiating factor. More recent experiments have documented an own-age bias (OAB), or a superior ability to recognize faces of one’s own age group when compared to faces of other age groups (Rhodes & Anastasi, 2012). Rhodes and Anastasi (2012) used a paradigm similar to that of the ORB studies mentioned above, with slight alterations of stimuli and presentation times. Both children and older adults acted as participants who studied a set of pictures of individuals representing several age groups. Participants were then asked to make a judgment regarding their perceived age. Subsequent recognition supported the notion of in-group bias in face recognition. Individuals had more accurate recognition for their own age group. Physiological evidence suggests that there is a specific difference in scalp electrical activity when viewing same-age individuals versus individuals of different age groups (Wiese, 2012). Differences based on group membership were found in early and late old/new effects between 300 and 800 milliseconds. These findings provide striking evidence demonstrating in-group biases in behavior and brain activity based on age and race.

Modulation of In-Group Bias

Evidence for the age- and race-based in-group biases indicate that perception of others is heavily dependent on our individual, biological characteristics. However, there are research groups who have taken a different approach to evaluate why certain faces are more difficult or easier to recognize. Sangrigoli, Pallier, Argenti, Ventureyra, and de Schonen (2005) recruited participants who belonged to one of three groups: Caucasian residents of France, native Koreans who had been residing in France for several months to over a decade, and Korean-born individuals who were adopted and raised by Caucasian families. The unexpected finding was that Korean individuals who were adopted by Caucasian families performed more accurately on recognition tasks for Caucasian, but not Asian, faces. Indeed, this finding lends support to the idea that experience is the driving factor behind our acquisition of an in-group and the associated increase in recognition accuracy.

Chiroro, Tredoux, Radaelli, and Meissner (2008) supported this view of categorization. They too present findings that suggest recognition bias is driven by more than just race and age. Chiroro and colleagues achieve this by recruiting a unique study population. Researchers recruited Black and White South African participants who studied Black and White faces of individuals from the United States and South Africa. Results demonstrated that participants exhibited an ORB for individuals from their own continent (Africa) but not for individuals from a different continent (North America). The finding that continent-congruence produced a drastic difference in accuracy offered further support for the idea that biological characteristics do not dictate social categorizations alone. Furthermore, this finding can be explained by an expertise model of face recognition, which suggests that experience with certain groups of faces is the driving force for superior face recognition (Tanaka & Gauthier, 1997).

Research indicates that broad social categorization, not only expertise, can lead to more accurate recognition. For example, a critical study by Hehman, Mania, and Gaertner (2010) asked participants to view faces presented in different formats. Faces were either grouped by university affiliation or by race. There was a bias in recognition accuracy in favor of own-race faces in the condition where faces were grouped by race. Recognition performance was superior for faces grouped as attending one’s own university versus another university in the university affiliation condition. This university affiliation difference was found despite the experiment using own- and other-race faces. Similarly, Bernstein, Young, and Hugenberg (2007) classified stimuli (faces) as part of one's own university or another university. Results demonstrated better recognition for faces classified as being a part of one’s university compared to those classified as attending another university. Another important aspect of the findings is that there were no perceptual differences between faces presented; the primary manipulations were category labels.

These are just a few examples of the different social categorization techniques researchers have used, which are based on pre-existing social groups and have resulted in facial recognition biases.

Experimentally-Defined Social Groups and Perceptually Equivalent Stimuli

Another sub-field of in-group bias research examined whether or not group biases can exert influence over perceptually equated stimuli (faces). Several methods established groups experimentally—based on university affiliation labels, personality traits, etc.—and asked participants to engage in recognition tasks with common experimental elements, such as equal number of faces of different races, all racially ambiguous faces, or all Caucasian faces as stimuli (Bernstein et al., 2007; Pauker, Weisbuch, Ambady, Sommers, & Ivcevic, 2009; Young & Hugenberg, 2010).

Social psychologists were the first to study experimentally created groups and how they affect participant behavior and judgments. Henri Tajfel and other early research groups established a paradigm in which participants were experimentally selected into a social group and asked to make decisions regarding behavior toward individuals within or outside that group (Diehl, 1990; Tajfel, Billig, Bundy & Flament, 1971). In order to establish groups that did not exist prior to the experiment, Tajfel et al. (1971) used a faux numerical estimation test whereby participants judged how many dots were present in a series of 40 clusters. They were told that they would be divided based on their estimation style into overestimators and underestimators. Participants subsequently favored in-group members when allotting different rewards and penalties. This paradigm of numerical estimation has also been utilized in facial recognition research (Ratner & Amodio, 2013a). Ratner and Amodio (2013) assigned participants to a group based on numerical estimation style. A face categorization task was used and participants categorized faces as overestimators or underestimators in a series of 60 trials. Researchers used six Caucasian male faces as stimuli for the physiological aspect of the study and recorded electrical scalp activity. Results demonstrated no reliable difference between accuracy of recognition, which they suggested was due to the simple nature of the task. However, they found a significantly greater amplitude for the N170 component in response to in-group faces compared to out-group. These researchers were some of the first to demonstrate early processing differences between groups resulting from a minimal group paradigm.

Pinter and Greenwald (2011) analyzed several techniques used to create these 'minimal groups' that were derived from Tajfel's early discovery. For example, researchers have often used faux personality quizzes and art ratings to separate individuals into groups and induce group connections. Participants who rated paintings were separated into 'red' or 'green' groups, which they were told groups were based on their artistic preferences. Pinter and Greenwald (2011) found that the simple act of studying names related to one's in-group could produce a robust minimal group categorization. This group categorization procedure involves telling participants that they are members of the 'red' or 'green' group and then asking them to study a series of names belonging to their group members. These methods have also been used within the face recognition literature. For example, Bernstein et al. (2007) established minimal groups using faux personality tests and red-green categorizations. This manipulation produced an effect similar to those created by the ORB, such that participants were more accurate at recognizing individuals of the in-group compared to the out-group. Bernstein et al. (2007) findings revealed that *experimentally* assigning participants to a group resulted in more accurate recognition. Collectively, these findings suggest that group effects, or biases toward the in-group, can be found in facial recognition even when there were no characteristic perceptual differences in the face stimuli.

Theoretical Roots

The research that examines experimentally-defined groups helps to compare a perceptual expertise model to a social cognitive model of face recognition. The use of minimal group procedures and perceptually equated stimuli allowed researchers to isolate the potential role of socio-cognitive factors in face recognition in the present research. The perceptual expertise model suggests that we hone our skills for particular categories of people based on experience with those groups (Rhodes, Lie, Ewing, Evangelista, & Tanaka, 2010). In this learning and skill acquisition process, individuals become better at extracting important features, or "invariant cues," for in-group members (Meissner & Brigham, 2001, Sporer, 2001). For example, individuals could learn that eyes are the most salient feature in the recognition process. These same individuals would become more skilled at differentiating between the eyes, and therefore faces, of in-group members due heightened experience and knowledge of the diagnostic nature of eyes in person identification.

A focus on features is also found in face-space models of face recognition. In this model, faces are stored within a framework that codes for features and combinations of features, (Meissner & Brigham, 2001). This framework suggests that experience with faces leads to the development of face exemplars. New faces are then compared to these exemplars and the degree of match can help or hinder later recognition. The important aspect of this model is that the exemplars within the face-space are dependent on experience. Therefore, more experience with in-group members would lead to better recognition of in-group faces, because of the greater degree of match with the exemplar. Under these perceptual feature-based perspectives, one would not expect to find any group differences between in-groups and out-groups when using a minimal group procedure. Minimal groups eliminate experience with a group by eliminating characteristic perceptual features of group members.

The social cognitive model suggests that group distinctions and categorizations alter the way in which people process faces (Hugenberg, Young, Bernstein, & Sacco, 2010). For example, the model suggests that people are motivated by group status to individuate in-group faces (i.e., process faces as distinct individuals). Similarly, people are motivated to process out-group faces in a categorical fashion (i.e., process faces as being members of a broader category or group). Social dynamics act on perceptual and behavioral outcomes by changing motivation. Researchers have suggested that this social cognitive bias would be borne by implicit value of in-group members. This value might be based on access to shared resources and collective gains (Sporer, 2001). The social cognitive theory would predict differences between in- and out-group faces on a facial recognition task using perceptually matched stimuli. This theory is not predicated on perceptual differences and experience with the individuals. The present experiments shed light on different factors contributing to in-group bias and give insights into their weight and influence on behavioral results.

Current Experiment

This series of experiments combined methodologies from cognitive psychology and social psychology to critically examine biases in facial recognition. First, minimal group paradigms were employed to create in-groups and out-groups. Individuals in the name-study experimental condition were asked to study in-group member names and individuals in the numerical estimation condition were asked to estimate the number of dots presented in a series of 10 images. Participants in the numerical estimation group were then designated as an overestimator or underestimator. All participants in both of the experimental conditions were then assigned a group color. Experiment 1 attempted to find group differences similar to studies examining the ORB and the OAB. Researchers assigned participants to groups and asked them to study faces of putative in-group and out-group individuals before a recognition test. Faces presented during the test in Experiment 2 were presented on a colored background similar to the study phase. The structure of the test phase allowed a more accurate comparison of recognition abilities between in-group and out-group faces that were seen during test. Experiment 3 looked at memory effects by extending the presentation time and reducing the number of faces presented to the participants to commit to memory from 32 to 20. This was used to assess whether the pattern of group differences was driven by memory. It was hypothesized that minimal group procedures would cause recognition accuracy, quantified through signal detection measures, to be superior for in-group faces. Additionally, reaction times were expected to be shorter and confidence ratings were expected to be higher in response to in-group faces compared to out-group faces.

GENERAL METHODS

Participants

Undergraduate students from Colorado State University participated for course credit or extra credit.

Materials

Images were selected from a database collected at Notre Dame University and were presented on a computer screen (Phillips et al., 2012). E-prime 2.0 was used for stimulus presentation and response collection. The subset of the stimuli included images of Caucasian individuals—both female and male. The stimuli included color, hair and clothing information. The images were taken in real-world settings, including in hallways and outdoors. Equal amounts of male and female faces were used for each experiment. Participants were seated in front of the computer screen and responded using a keypad.

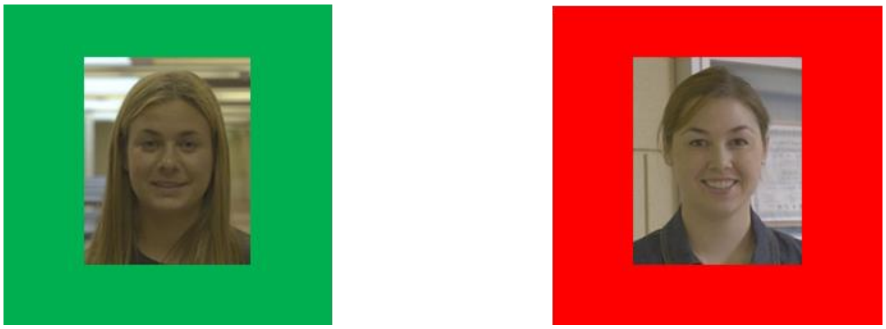

Figure 1. Example stimuli being presented on red and green backgrounds. Faces were presented sequentially and individually.

Design

Independent variables included a between-subjects factor of type of minimal group induction (including a control condition) and a within-subjects factor of group membership (in-group or out-group). Recognition performance, reaction time, and confidence ratings were measured in response to test faces as dependent measures.

Procedure

Each participant was randomly assigned to a minimal group or control condition after providing consent to participate. The experiment involved an intentional encoding task where participants were instructed to do their best to remember a series of faces. Participants were asked to report as many of the states within the United States for two minutes as a filler task and then asked to complete a recognition test.

Calculations for d' (sensitivity) and C (response bias) were made based on the standard parametric signal detection model. Extreme values on hits and false alarms (based on proportions) were corrected by subtracting or adding 0.01 respectively. For example, scores of 1 on proportions of hits and/or false alarms would be corrected to 0.99 and scores of 0 would be corrected to 0.01. This adjustment allows for more reliable calculations of sensitivity and response bias. D-prime scores were calculated by subtracting false alarm rates from hit rates. This calculation was conducted using normalized scores. Scores on C were calculated by first subtracting the half of the d' score from the normalized false alarm rate. This value was then multiplied by -1. Positive values indicated a bias toward "new" and higher positive values indicated more conservative biases. The reverse was true for negative values. Responses that were not "old" or "new" were excluded from analysis. Reaction time—latency to an old/new response—was measured in response to each test phase stimulus in milliseconds by the E-prime program. Participants also made confidence ratings regarding their old/new decisions and reported them using a scale from 0% to 100% via the computer program. Scores that exceeded 100 were excluded from analysis.

EXPERIMENT 1

This experiment compared the effectiveness of two types of minimal group procedures to a control condition in establishing in-group bias. During the study phase, faces were presented on red or green backgrounds. An equal number of red and green backgrounds were randomly assigned to the images. The test phase in this experiment did not include the colors that identified faces as either in-group or out-group. This color information was only included during the study phase to provide a more ecologically valid test of recognition.

Importantly, stimuli were perceptually equated and were randomly placed into either in-group or out-group and paired with a colored background. This would suggest that any differences found between in- and out-group faces would be due to social cognitive factors and not perceptual expertise.

Hypotheses

A larger percentage of in-group faces were expected to be accurately recognized as ‘old’—or seen before—in the test phase compared to out-group faces. Reaction time was hypothesized to be shorter for in-group faces and confidence ratings were expected to be higher in response to ratings of in-group faces.

Method

Undergraduate students (N = 126) participated in Experiment 1 for course credit or extra credit (M = 19.74, SD = 1.96; Female = 92). Equal numbers of participants were included in each condition, control, name-study and numerical estimation (N = 42 per group).

Materials

Thirty-two images were used for the study phase and 64 faces were used for the test phase. Colored backgrounds were not used during the test phase.

Design

This experiment had a mixed design with a between-subjects factor of minimal group procedure (including a control condition) and a within-subjects factor of group membership (in-group or out-group). The within-subjects factor was only present in the experimental conditions, where an in- and out-group were established. Percent correct, reaction time, and confidence ratings were measured for both in-group and out-group faces as dependent measures.

Procedure

Each participant was randomly assigned to the name-study, numerical estimation or control condition after providing consent to participate. Subsequent experiments utilized only the most effective minimal group condition based on recognition accuracy (name-study), and the control condition. In the name-study condition, each participant was be given 2 minutes to study a list of 16 names of individuals who, the participants were told, were in their in-group (Pinter & Greenwald, 2011). This group was designated by random assignment to either the red or the green group.

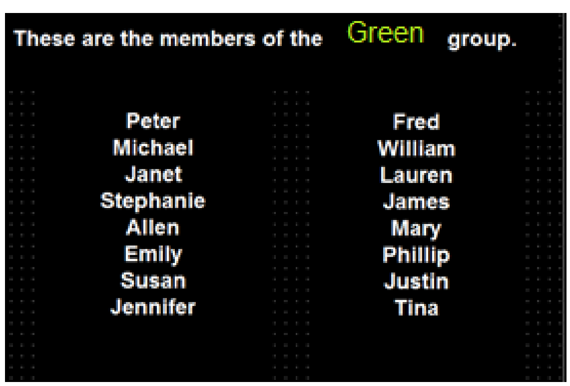

Figure 2. Name-study condition example stimuli. Participants were presented with in-group names.

The numerical estimation condition asked participants to estimate the number of dots in a series of 10 clusters. They were then designated to a red or green group, which they were told was based on their style of estimation—overestimator or underestimator. The control condition did not receive any additional instruction or group assignment prior to the study phase.

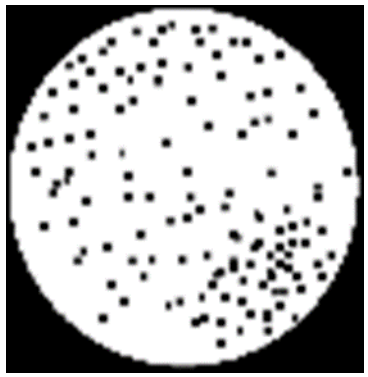

Figure 3. Numerical estimation example stimuli. Participants were asked to estimate the number of dots presented in formations such as this.

Participants studied 32 faces (16 female, 16 male) for five seconds each. A filler task was then presented and participants spent two minutes recording state names. A 64-item recognition test was then administered. Test faces were presented randomly to participants who were asked to identify the faces as either 'old' or 'new.' A response of 'old' indicated having seen the face in the study condition and a 'new' response indicated not having seen the face before. New images of individuals from the study phase were used during test and randomly interspersed with entirely new face stimuli. Reaction times were recorded and participants rated their confidence in their response following each old/new classification.

Results

Number of hits and false alarms, discrimination and response criterion were calculated for and compared between each experimental group. C and d' scores were used to indicate accuracy and response bias. Reaction time and confidence ratings were compared between experimental groups as well as in between in-group and out-group comparisons. Percent correct scores were recorded as a measure of accuracy for in-group and out-group comparisons. These scores were calculated using the number of 'old' items that were categorized as old (or seen before) by the participants out of the 16 possible old faces for both in- and out-groups. One-way ANOVAs were used to analyze measures between the three conditions. Three 2x2 mixed factors ANOVA were used to compare percent correct, reaction time and percent correct for in-group versus out-group faces (within-subjects factor) between groups (between-subjects factor).

Recognition Accuracy

No significant differences were found using one-way ANOVAs between the three experimental groups on hits, $F(2,123) = 0.692$, $p = .503$, $\eta_p^2 = .011$, false alarms, $F(2,123) = 0.135$, $p = .874$, $\eta_p^2 = .002$, d', $F(2,123) = 0.934$, $p = .396$, $\eta_p^2 = .015$, and C, $F(2,123) = 0.267$, $p = .766$, $\eta_p^2 = .004$. This can be viewed in Table 1.

Table 1

Experiment 1 Means for Signal Detection Measures Between Experimental Groups

	Name Study	Numerical Estimation	Control
Hits	0.47 (0.17)	0.45 (0.15)	0.49 (0.15)
False Alarms	0.27 (0.16)	0.27 (0.13)	0.28 (0.14)
d'	0.64 (0.37)	0.53 (0.36)	0.62 (0.40)
C	0.41 (0.49)	0.41 (0.43)	0.35 (0.44)

Results from a 2x2 mixed factorial ANOVA for percent correct—a measure of hits—demonstrated no within-subject effect of group membership, $F(1,82) = 0.442$, $p = .508$, $\eta_p^2 = .005$ (Figure 1). No effects were found for the between-subjects effects of condition, $F(1,82) = 0.442$, $p = .508$, $\eta_p^2 = 0.05$, or the interaction between group membership and condition, $F(1,82) = 0.297$, $p = .588$, $\eta_p^2 = .004$.

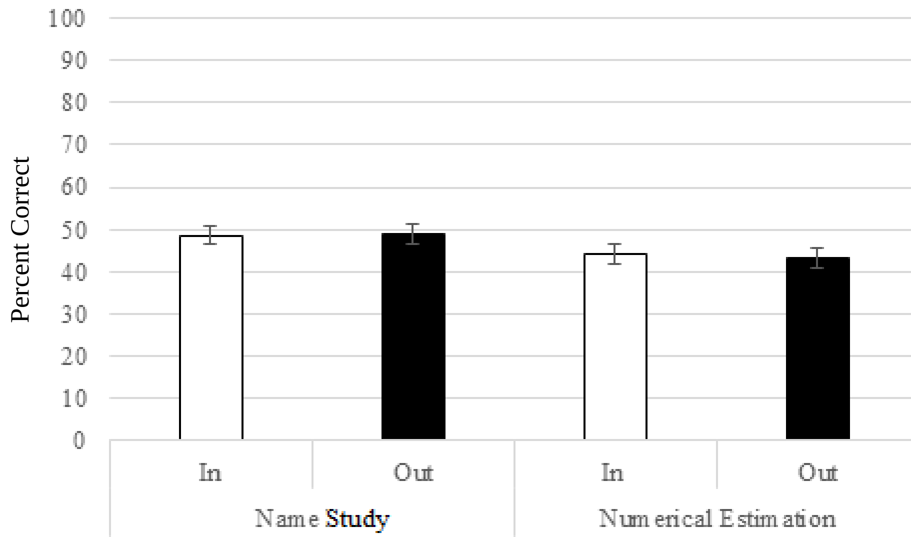

Figure 4. Mean percent correct for in- and out-group faces in the two experimental conditions.

Reaction Time

There were no significant difference based on reaction time between the three experimental groups, $F(2,123) = 0.383$, $p = .682$, $\eta_p^2 = .006$. A 2x2 mixed factorial ANOVA demonstrated that the within-subjects factor of group membership did not have an effect on reaction times, $F(1,82) = 0.231$, $p = .632$, $\eta_p^2 = .003$ (Figure 2). Similarly, no between-subjects effect of condition, $F(1,82) = 0.011$, $p = .918$, $\eta_p^2 = .000$, nor an interaction was found, $F(1,82) = 0.879$, $p = .351$, $\eta_p^2 = .011$. Table 2 provides reaction time values for Experiment 1.

Table 2

Experiment 1 Average Reaction Time in Milliseconds

	Name Study	Numerical Estimation	Control	Overall
Overall	3368.57 (984.67)	3417.13 (858.79)	3245.86 (873.17)	3343.85 (902.91)
Old	3436.26 (1045.90)	3455 (868.03)	3436.26 (948.17)	3403.97 (950.99)
New	3301.49 (1003.35)	3379.06 (944.44)	3171.37 (903.25)	3283.97 (947.51)
In-group	3365.86 (1201.68)	3476.77 (924.59)		
Out-group	3504.14 (1142.24)	3433.23 (917.60)		

Confidence Ratings

There were no significant difference based on confidence ratings between the three experimental groups, $F(2,123) = 0.380$, $p = .685$, $\eta_p^2 = .006$. Results from a 2x2 mixed factorial ANOVA are illustrated in the confidence ratings reported in Table 3. There were no effects based on group membership, $F(1,82) = 0.224$, $p = .637$, $\eta_p^2 = .003$, no effect of experimental group, $F(1,82) = 0.24$, $p = .626$, $\eta_p^2 = .003$, and no interaction between group membership and condition, $F(1,82) = 0.029$, $p = .865$, $\eta_p^2 = .000$.

Table 3

Experiment 1 Average Confidence Ratings (0-100%)

	Name Study	Numerical Estimation	Control	Overall
Overall	71.99 (13.96)	70.47 (11.59)	69.34 (13.40)	70.60 (12.96)
Old	72.86 (13.78)	71.21 (11.35)	70.23 (13.10)	71.43 (12.72)
New	71.11 (15.00)	69.74 (12.57)	68.46 (14.27)	69.77 (13.91)
In-group	73.40 (15.06)	71.49 (12.46)		
Out-group	72.33 (13.78)	70.92 (10.99)		

Summary and Statement of the Problem

The hypotheses for Experiment 1 suggested that an in-group bias could be detected in face recognition task through experimentally manipulating groups. Performance on all measures did not vary significantly between experimental conditions (control, name study, numerical estimation) or based on group membership (in- and out-groups). The results of this experiment do not support the idea that minimal groups can impact face recognition.

Limitations

This experiment was limited by test phase format. As it was, participants' performance could only be judged based on percent correct, instead of a full signal detection analysis. Showing test phase faces on red and green backgrounds—similar to the study phase—would allow for additional comparisons, namely the degree of sensitivity and discrimination between in- and out-group faces. In addition, accuracy was low overall across participants. The average number of hits was around 15 (out of 32 possible) and did not differ significantly between groups. It is possible that increasing performance on the task could inform the question of whether or not minimal groups create meaningful differences on face perception.

EXPERIMENT 2

Experiment 2 built on the first experiment by altering the test phase. This was done to allow for signal detection comparisons between in-group and out-group faces. In Experiment 1, faces were presented in isolation during the test phase (i.e., no colored background was used). In comparison, Experiment 2 presented faces on red or green backgrounds at test. “Old” faces that were seen during the study phase were presented on the same color backgrounds later in the test phase. “New” faces were presented on quasi-random red and green backgrounds (an equal number of red and green backgrounds were used). Signal detection analysis techniques were then used to examine response bias and discrimination between in- and out-group faces by using this type of test phase. In essence, Experiment 2 was used to more deeply examine recognition behavior of participants and serve as a partial replication of Experiment 1.

Hypotheses

Similar to Experiment 1, reaction times were hypothesized to be shorter for in-group faces and confidence ratings were expected to be higher for in-group faces. Recognition accuracy was expected to demonstrate greater sensitivity and a more conservative response bias toward in-group faces.

Method

Experiment 2 included 104 individuals who were recruited using the participant pool at Colorado State University and received course credit for participation. Participants had a mean age of 19.85 (1.98) years with 46 females and 58 males.

Materials

The stimuli used for Experiment 2 were the same as those used in Experiment 1. Thirty-two faces were presented during the study phase and 64 were presented at test—half of the faces were foils and had not been seen prior to the test phase. Faces were presented on red and green backgrounds during both the study and test phases.

Design

Experiment 2 had a mixed design with a between-subjects factor of minimal group procedure (including a control condition) and a within-subjects factor of group membership (in-group or out-group).

Procedure

Experiment 2 was identical to Experiment 1 with a few exceptions. In Experiment 2, faces during test were also presented on red or green backgrounds. Faces seen in the study phase were presented on the same background color and new faces were assigned red or green, pseudo-randomly—with red and green backgrounds occurring the same number of times.

Results

Hits, false alarms, d', and C were calculated using signal detection theory. Additionally, reaction time and confidence ratings were measured for both in-group and out-group faces as dependent measures. Data analysis included an additional comparison of C and d' values between in-group and out-group recognition scores that were impossible to compute in Experiment 1. Mixed factorial (2x2) ANOVAs were used to analyze the data for each measure, with experimental group as the between-subjects factor and group membership as the between-subjects factor.

First analyses looked at old/new comparisons. No significant differences on reaction time were found between groups, $F(1,102) = 0.288$, $p = .593$, $\eta_p^2 = .003$, between old and new stimuli, $F(1,102) = 3.425$, $p = .067$, $\eta_p^2 = .032$, nor in the interaction effect, $F(1,102) = 0.518$, $p = .473$, $\eta_p^2 = .005$. Confidence ratings also produced no significant within-subject effect of old versus new, $F(1,102) = 5.319$, $p = .023$, $\eta_p^2 = .050$, no between-subjects effect of group, $F(1,102) = 0.473$, $p = .493$, $\eta_p^2 = .005$, and no interaction effect, $F(1,102) = 0.135$, $p = .714$, $\eta_p^2 = .001$.

Recognition Accuracy

No main effect of group on accuracy, $F(1,102) = 0.001$, $p = .976$, $\eta_p^2 = .000$, and no interaction effect between group and old versus new on accuracy, $F(1,102) = 1.392$, $p = .241$, $\eta_p^2 = .013$ were found. However, there was a significant within-subjects difference between hits ($M = 0.46$, $SD = 0.17$) and false alarms ($M = 0.30$, $SD = 0.16$), $F(1,102) = 153.241$, $p < .001$, $\eta_p^2 = .600$ (Table 4). This suggests that the recognition task was effective, and participants more often labeled “old” faces as “old” instead of “new.” However, the level of recognition overall was still very low.

Table 4

Experiment 2 Signal Detection Measures

	Hits	False Alarms	d'	C
Name Study	0.47 (0.17)	0.29 (0.16)	0.54 (0.43)	0.37 (0.50)
Control	0.45 (0.16)	0.31 (0.17)	0.44 (0.42)	0.37 (0.48)
Overall	0.46 (0.17)	0.30 (0.16)	0.49 (0.43)	0.37 (0.49)
In-group	0.45 (0.18)	0.28 (0.19)	0.54 (0.59)	0.41 (0.56)
Out-group	0.48 (0.19)	0.29 (0.17)	0.57 (0.57)	0.36 (0.52)
Green	0.45 (0.19)	0.31 (0.19)	0.46 (0.59)	0.39 (0.57)
Red	0.45 (0.17)	0.31 (0.19)	0.45 (0.58)	0.38 (0.51)

The results of a 2x2 mixed factorial ANOVA suggest that there are no significant differences based on group membership, $F(1, 102) = 0.026$, $p = .871$, $\eta_p^2 = .000$, experimental group, $F(1,102) = 1.281$, $p = .260$, $\eta_p^2 = .012$, and no interaction effect on d', $F(1,102) = 0.064$, $p = .800$, $\eta_p^2 = .001$.

A 2x2 mixed factorial ANOVA demonstrated no effects for criterion, $F(1,102) = 0.005$, $p = .942$, $\eta_p^2 = .000$. The mean criterion for the name-study condition was equally conservative compared to the control condition. This can be viewed in Figure 3. In the control condition, green faces ($M = 0.39$, $SD = 0.57$) demonstrated a nearly identical response bias compared to red faces ($M = 0.38$, $SD = 0.51$). In the name study condition, in-group faces ($M = 0.41$, $SD = 0.56$) were more conservative compared to out-group faces ($M = 0.36$, $SD = 0.52$), but not dramatically so. No significant differences were found for group membership, $F(1,102) = 0.820$, $p = 0.367$, $\eta_p^2 = .008$, or an interaction effect, $F(1,102) = 0.021$, $p = 0.573$, $\eta_p^2 = .003$.

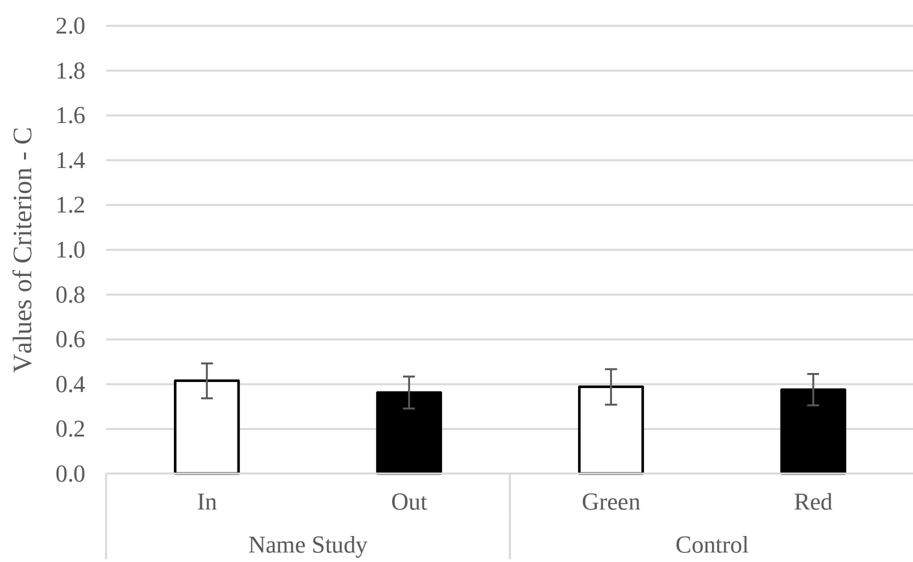

Figure 5. Experiment 2: Comparison of Response Bias between Experimental Groups.

Reaction Time

No significant differences were found on measures of reaction time. An old/new effect was marginally significant, $F(1,102) = 3.425$, $p = .067$, $\eta_p^2 = .032$, such that old faces (M = 3348.74, SD = 1053.66) took longer to respond to compared to new faces (M = 3226.58, SD = 1044.96). There was no main effect of experimental group, $F(1,102) = 0.288$, $p = .593$, $\eta_p^2 = .003$, based on a 2x2 mixed factorial ANOVA. Likewise, there was no within-subjects effect of group membership, $F(1,102) = 0.093$, $p = .761$, $\eta_p^2 = .001$, nor an interaction effect, $F(1,102) = 0.406$, $p = .525$, $\eta_p^2 = .004$.

Confidence Ratings

An old/new effect was the only significant difference in confidence ratings, $F(1,102) = 5.319$, $p = .023$, $\eta_p^2 = .050$. Subjects were more confident for their responses to old faces (M = 67.69, SD = 13.70) compared to new faces (M = 69.05, SD = 12.73). A 2x2 mixed factorial ANOVA demonstrated no main effects of group membership, $F(1,102) = 0.016$, $p = .901$, $\eta_p^2 = .000$, or experimental group, $F(1,102) = 0.476$, $p = .492$, $\eta_p^2 = .005$. Additionally, no interaction was found, $F(1,102) = 0.154$, $p = .696$, $\eta_p^2 = .002$.

Discussion

Experiment 2 replicated and built on the first experiment by addressing the question of whether or not sensitivity (d’) and response bias (C) are impacted by a simple minimal group procedure. There were no significant differences between groups. These results suggest that accuracy is not impacted by minimal group procedures.

Limitations

The major limitation of this study was poor performance overall on the recognition task. The task of remembering 32 faces and recognizing them within the same number of foils proved challenging for the participants. Lack of accuracy or task difficulty may have overshadowed smaller group membership differences. These findings inspired a critical examination the validity of their dependent variables in the context of the experimental task.

EXPERIMENT 3

The design of Experiment 3 directly addressed results from the first two experiments in this series. Recognition accuracy was low across groups in Experiment 1 and Experiment 2. Researchers attributed this to the number of faces to be recalled and the duration of stimulus presentation, 64 faces and 5 seconds respectively. It was hypothesized that potential biases in recognition behavior were overshadowed by poor memory performance in the first two experiments. Experiment 3 reduced the number of to-be-remembered stimuli from 32 to 20 and increased the duration of presentation to 10 seconds per stimulus. Previous research suggests shortening list length and increasing presentation time directly improve encoding and later recognition performance (Gronlund & Elam, 1994; Reynolds & Pezdek, 1992; Strong, 1912).

Hypotheses

The hypotheses for Experiment 3 are the same as stated for Experiment 2. Differences between in-group and out-groups are expected on each dependent measure. Improved accuracy, shorter reaction time, and greater levels of confidence are expected for in-group faces.

Method

Experiment 3 included 92 individuals who were recruited using the participant pool at Colorado State University and received course credit for participation. Participants had a mean age of 18.95 (1.22) years with 39 males and 53 females.

Materials

The stimuli used for Experiment 3 matched those of the first two experiments. Twenty faces were presented during the study phase and 40 were used during the test phase—20 of the same faces used at study and 20 foils. Faces were presented on red and green backgrounds during both the study and test phases, with equal numbers of red and green backgrounds.

Design

The design of Experiment 3 was no different from that of Experiment 2. Researchers utilized a 2x2 mixed design minimal group procedure (between-subjects) and group membership (within-subjects) as variables of interest. Accuracy (using signal detection), reaction time and confidence ratings served as dependent variables.

Procedure

Participants studied 20 faces (10 female, 10 male) for ten seconds per image. This is a marked reduction in the number of images to be remembered and is a doubling of the presentation time compared to the first and second experiments. All other aspects of the experiment procedure were the same as Experiment 2.

Results

The analyses conducted for Experiment 3 were the same used for Experiment 2 and are detailed previously.

Recognition Accuracy

A mixed factorial ANOVA demonstrated a significant main effect of minimal group procedure on false alarms, $F(1,90) = 5.841$, $p = .018$, $\eta_p^2 = .061$. Within the control condition, in-group [green] faces (*M* = 0.22, *SD* = 0.18) had a greater amount of false alarms compared to the out-group [red] faces (*M* = 0.18, *SD* = 0.16). The in-group in the name study condition (*M* = 0.28, *SD* = 0.18) had the same number of false alarms as the out-group faces (*M* = 0.28, *SD* = 0.18). A main effect of group membership and an interaction were not found, $F(1,90) = 1.772$, $p = 0.186$, $\eta_p^2 = .019$ and $F(1,90) = 1.772$, $p = 0.186$, $\eta_p^2 = .019$ respectively.

A difference in the number of hits between name-study (In-group, *M* = 0.51, *SD* = 0.20; Out-group, *M* = 0.49, *SD* = 0.19) and control groups (In-group [green], *M* = 0.44, *SD* = 0.18; Out-group [red], *M* = 0.43, *SD* = 0.22) approached significance, $F(1,90) = 3.650$, $p = .059$, $\eta_p^2 = .039$. No main effect of group membership was found, $F(1,90) = 0.280$, $p = 0.598$, $\eta_p^2 = .003$, and no interaction effect was found, $F(1,90) = 0.187$, $p = 0.666$, $\eta_p^2 = .002$.

No effects were found based on the signal detection measure, d' (Table 5). No main effect of group membership was found, $F(1,90) = 0.280$, $p = .598$, $\eta_p^2 = .003$. Similarly, there was no main effect of experimental group, $F(1,90) = 2.118$, $p = 0.149$, $\eta_p^2 = .023$, and no interaction, $F(1,90) = 0.803$, $p = .186$, $\eta_p^2 = .019$.

Table 5

Experiment 3 Signal Detection Measures

	Hits	False Alarms	d'	C
Name Study	0.50 (0.16)^	0.28 (0.16)*	0.68 (0.51)	0.34 (0.44)*
Control	0.44 (0.17)^	0.20 (0.15)*	0.84 (0.53)	0.60 (0.53)*
Overall	0.47 (0.17)	0.24 (0.16)	0.76 (0.52)	0.47 (0.50)
In-group	0.51 (0.20)	0.28 (0.18)	0.75 (0.69)	0.34 (0.58)
Out-group	0.49 (0.19)	0.28 (0.18)	0.66 (0.73)	0.39 (0.55)
Green	0.44 (0.18)	0.22 (0.18)	0.78 (0.68)	0.58 (0.56)
Red	0.43 (0.22)	0.18 (0.16)	0.97 (0.80)	0.70 (0.63)

* = indicate significant differences between the two groups with an alpha of .05.

Differences were found on response criterion—C. A main effect of minimal group procedure between control (In-group [green], *M* = 0.58, *SD* = 0.56; Out-group [red], *M* = 0.70, *SD* = 0.63) and name-study (In-group, *M* = 0.34, *SD* = 0.58; Out-group, *M* = 0.39, *SD* = 0.55) groups was found on C, $F(1,90) = 6.384$, $p = .013$, $\eta_p^2 = .066$. This can be viewed in the following Figure 7. The differences on signal detection measures, especially on criterion and false alarms, suggest that differences in motivation and strategy exist between the name study and control group. There was not a main effect of group membership on response bias, $F(1,90) = 2.845$, $p = .095$, $\eta_p^2 = .031$. No interaction was found, $F(1,90) = 0.398$, $p = .530$, $\eta_p^2 = .004$.

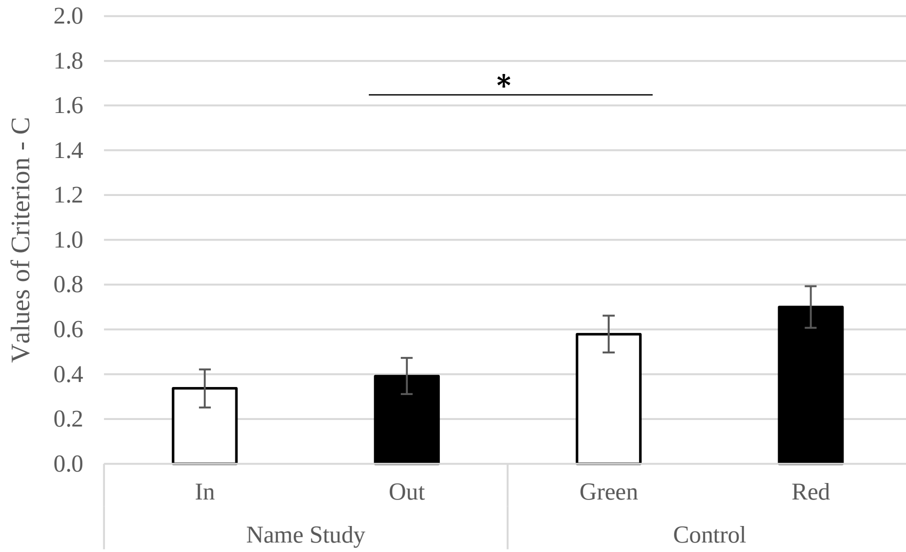

Figure 6. Experiment 3: Comparison of Response Bias between Experimental Groups. Asterisks indicate a significant difference with alpha set to .05.

Reaction Time

No significant differences in reaction time were found using a 2x2 mixed factorial ANOVA. No main effects of group membership or experimental group were found, $F(1,90) = 0.084$, $p = .772$, $\eta_p^2 = .001$ and $F(1,90) = 0.344$, $p = .559$, $\eta_p^2 = .004$ respectively. No interaction was found, $F(1,90) = 0.334$, $p = .565$, $\eta_p^2 = .004$.

Confidence Ratings

There were also no significant differences found for confidence ratings across the board. No main effect of group membership, $F(1,90) = 0.126$, $p = .723$, $\eta_p^2 = .001$, or experimental group, $F(1,90) = 0.149$, $p = .701$, $\eta_p^2 = .002$, were found. Similarly, there was no interaction, $F(1,90) = 0.930$, $p = .337$, $\eta_p^2 = .01$.

Summary and Statement of the Problem

The primary goal of Experiment 3 was to determine whether or not improving performance on the memory task, through attempting to enhance recognition of previously seen faces, would reveal differences between in- and out-groups. Instead of isolated differences between in-group and out-groups, the predominant group differences existed between experimental (name study) and control group. Significant differences in response bias, C, and false alarms suggest that minimal group procedures alter participants' implicit or explicit strategies. In other words, these results suggest that minimal group procedures modulate response bias. Moderate evidence suggests that minimal group procedures also alter discriminability of experimental stimuli.

Limitations

One limitation involved the analysis of control and experimental groups for in- and out-group comparisons. Control groups were not randomly assigned a minimal group and therefore researchers used red and green group means to compare to in- and out-group means. There may have been confounds within the comparisons. Performance was improved to nearly 50% accuracy on average, but this could be further enhanced to distinguish group differences.

GENERAL DISCUSSION

This series of experiments has synthesized several aspects of an important body of literature into a set of studies to evaluate the recognition of categorized individuals. These results provide minimal support for the social cognitive perspective of face recognition, because the presence of a minimal group influenced response bias despite no perceptual differences between the stimuli. This suggests that simple social categorization slightly alters how we recognize faces. However, no differences were found between in-groups and out-groups. This suggests that perceptual expertise drives biases in facial recognition. The results of Experiment 3 demonstrate how transient mechanisms of social categorization alone can act as a weak source of in-group bias in face recognition.

Recognition Performance

Participants in the control group adopted more conservative response biases toward all faces in Experiment and 3. These comparisons were significant for both experiments. This suggests that the way faces are approached and how they are recognized later is related to the imposition of social groups.

Experiment 3 was conducted to improve performance on the experimental task. Hits and false alarm values did not powerfully illustrate this expected improvement. Only the experimental group increased their hit scores from 45% to 50% accuracy. Despite the small improvement in accuracy, the change in procedure did impact response bias and false alarm group differences. As mentioned previously, these group differences suggest a different strategic approach used for the experimental group when compared to the control group.

The results of these experiments map on to other findings of research groups who have experimentally imposed social categorization. Similar to the present series of experiments, studies presented in the literature often use stimuli that are perceptually equated. Through equating stimuli, researchers are reasonably able to attribute any group differences to the manipulation of group status and rule out perceptual expertise. Hugenberg and colleagues have repeatedly demonstrated this effect of minimal groups on facial recognition. His research group utilizes colored backgrounds paired with university labels in order to assess the role of experimentally defined groups, as these stimuli are actually randomized and not distinct for any group (Bernstein et al., 2007). Additionally, this research group has demonstrated that in-group faces are more susceptible to disruption of configural processing compared to out-group faces (Hugenberg & Corneille, 2009). This finding suggests that in-group faces are processed differently compared to out-group faces—as a whole versus based on features, respectively. This effect mirrors that of own-race faces in the literature. Better expression recognition was also found for in-group faces and this effect was especially potent when the minimal group was established prior to encoding versus post-encoding (Hugenberg et al., 2010; Young & Hugenberg, 2010). The work from the Hugenberg lab generally demonstrates a small to moderate effect size (d = 0.38, 0.40, 0.46), for group membership differences (in-group vs. out-group; Bernstein et al., 2007, Young, Bernstein, & Hugenberg, 2010).

A similar group of studies has examined the role of establishing experimentally defined groups on facial recognition and abolishing the superior recognition for own-race faces. Hehman and colleagues have demonstrated robust in-group and out-group effects by simply changing the university label assigned to a face on a colored background (Hehman et al., 2010; Hehman, Stanley, Gaertner, & Simons, 2011). These studies have strengthened the behavioral evidence for in-group and out-group differences—on perceptually equated stimuli—in facial recognition. However, there is also evidence that counter these findings from studies using ambiguous race faces and neuroscience techniques in addition to behavioral methods (Pauker & Ambady, 2009; Ratner & Amodio, 2013; Van Bavel & Cunningham, 2012). This evidence will be presented and elucidated later in the discussion.

The significant results found in this series of experiments based largely on response bias are few compared to the number of non-significant comparisons found. Importantly, there were no significant differences between in-group and out-group stimuli. No main effects of group membership were found. Additionally, Experiment 1 and Experiment 2—to some degree—did not show the same significant effects of Experiment 3. While this could be attributed to improving the effectiveness of the experimental task in Experiment 3, this finding could also be attributed to a small effect of minimal group procedures on facial recognition overall. In fact, the literature presents examples of the inability of minimal group procedures to produce reliable, significant effects (Herzmann & Curran, 2013; Pauker & Ambady, 2009; Van Bavel & Cunningham, 2012). It is possible that the significant results found in the literature and in this study were and are due to sampling error and not meaningful differences between in- and out-groups.

Van Bavel and colleagues have been studying the effect of minimal group procedures on the brain and behavior. In his lab's typical paradigm, participants are split into arbitrary groups labeled the leopards and the tigers for the duration of the experiment (Van Bavel, Packer, & Cunningham, 2008). Faces are then presented within the fMRI magnet and performance and activity is then measured. Researchers found differences in the fusiform gyrus, orbitofrontal cortex, amygdala and portions in the striatum between presentations of in- and out-group faces.

This effect was not generally supported by the behavioral data examining recognition. No memory effects were found despite finding differences in activation of the FFA for in-group and out-group recognition (van Bavel, Packer, & Cunningham, 2011). In a follow-up study, researchers were only able to find a small effect of group membership on recognition memory performance.

Similarly, Herzmann and Curran (2013) found differences in event-related potential components—between 300-500ms post-stimulus and 600-900ms post-stimulus—for group membership comparisons. Again, however, these results did not appear in behavioral effects. Ratner and Amodio (2013) found no significant differences in recognition accuracy between in-group and out-group photos. They did, however, find a significantly greater amplitude for the N170 component for in-group faces over out-group faces. This lack of consistency between measures and the magnitude of behavioral effects found suggest that this effect is either small or not present at all.

A final vein of research that is relevant to the current studies involve the study of ambiguous race faces. These studies used perceptually equated stimuli (faces of mixed race often created using computer software) and altered the information provided about the faces during study. Black and White participants were recruited and the ambiguous race faces were labeled as being "White" or "Black," thereby establishing in-group and out-group stimuli. Mixed results have been found throughout this literature. Some researchers find that ambiguous faces that are labeled "White" are better remembered than those labeled "Black" (Pauker & Ambady, 2009). However, these effects have been found to be non-significant elsewhere (Hourihan, Fraundorf, & Benjamin, 2013). It is clear that there is contradictory evidence and uncertainty regarding the ability of experimentally created social groups to create meaningful differences in the literature.

This uncertainty paired with the publication bias toward significant differences over null results suggests that researchers should be critical of this effect. The presence and potential size of this effect—creating in-group bias for equated stimuli—can be better understood though a synthesis of all relevant studies and their effect sizes.

Reaction Time

The majority of comparisons analyzing reaction time throughout these experiments presented no significant differences. Experiment 2 presented an old/new effect in reaction time. Old faces had longer reaction times compared to new faces. These findings go against the results of Ratner and Amodio (2013) in their study examining minimal group differences where reaction time and event-related potentials were dependent measures. These researchers found a significant difference between in- and out-groups on log-transformed reaction times. This group membership difference was not found in the current research. In their experiment, van Bavel et al. (2008) asked participants to categorize faces by race or by team (minimal groups). Within the team categorization task, in-group faces were more quickly recognized compared to out-group faces. Again, this does not parallel the current findings. A potential explanation for the disparity between the literature and the current research is that minimal group assignment, in the present studies, did not create meaningful enough differences between in-groups and out-groups. Therefore, reaction times did not differ. An alternative justification is that these differences in reaction time are due to publication bias that does not often report null results.

Confidence Ratings

Measures of confidence are sparse in experiments using minimal group procedures and perceptually-equated stimuli. However, confidence ratings were and are quite common in the broader field studying in-group bias in facial recognition. The typical finding is that participants are both more accurate *and* more confident in their recognition of in-group faces (Chiroro & Valentine, 1995). For example, a Black individual would be more likely to recognize another black individual and then rate themselves as more confident in their response. Our studies were hypothesized to show a similar relationship. Recognition of in-group faces, established using a minimal group procedure, was expected to result in higher ratings of confidence. This was not the case. Evidence that could explain the lack of significant differences can be seen in research on eye-witness testimony and behavior. This research suggests that there is a weak association between accuracy and confidence in judgment (Meissner, Brigham & Butz, 2005). This weak association is generally found in response to out-group, or other-race, faces and individuals. This suggests that if minimal group procedures were more effective, then we would see this differential effect in the diagnostic nature of confidence ratings in response to in- and out-group members.

Conclusions

The findings of these experiments suggest experimentally-defined social groups do not have a strong influence on face recognition. Lack of significant differences, in sensitivity especially, between in-group and out-group faces demonstrate the manipulations' inability to establish meaningful group membership. The aim of this project was to try and tease apart the complicated problem of in-group bias, or preferential treatment towards own-group members. This series of experiments appears to have oversimplified the bias and eliminated group differences in doing so. While simply categorizing faces as in- and out-group may play a small role in driving inter-group bias, it is clearly not a primary component. These results suggest that either experience with group members, inherent perceptual differences between stereotypical group members, or some combination of the two, are the primary determinants of in-group bias. Results suggest little, if any, support for the social cognitive theory.

Overall Limitations

The size of the effect of interest was the primary limitation of this study. The effect of minimal groups, when found, is generally small when looking across studies that use perceptually equated stimuli. Small effects can go unnoticed even if they are reliable and the variance in outcomes is much greater with small effects and smaller sample sizes. A meta-analysis of the effect of minimal group procedures on effect size could greatly aid in answering the question of whether or not minimal groups are effective. In doing so, insights into the components of in-group bias could be better explained. Secondarily, performance on the recognition task could be improved further, despite the adjustments made in Experiment 3. Ability to carry out the experimental task is crucial to the validity of the results calculated.

Future Directions

As mentioned above, a synthesis of multiple experiments with small sample sizes is integral to the forward trajectory of the study of minimal groups. Without doing so, determining whether the small effect exists at all is difficult. Similarly, additional comparisons of minimal group procedures must be conducted. The many ways of creating experimentally defined groups introduce extraneous variables that are largely unaccounted for in the discussion of results and conclusions. Researchers should confirm that the constructs being studied are indeed the same across research groups. Group simulations could also be enhanced with current techniques, such as simultaneous ERP and fMRI recordings. These techniques have already provided insights in other realms of in-group bias, including the ORB and the OAB. They can be used to tease apart meaningful effects from spurious.

REFERENCES

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*(8), 706–712. doi:10.1111/j.1467-9280.2007.01964.x

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327.

Chiroro, P. M., Tredoux, C. G., Radaelli, S., & Meissner, C. A. (2008). Recognizing faces across continents: the effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin & Review*, *15*(6), 1089–1092. doi:10.3758/PBR.15.6.1089

Desteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science*, *15*(5), 319–324.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology. General*, *115*(2), 107–17. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2940312

Hehman, E., Mania, E. W., & Gaertner, S. L. (2010). Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect. *Journal of Experimental Social Psychology*, *46*(2), 445–448. doi:10.1016/j.jesp.2009.11.008

Hehman, E., Stanley, E. M., Gaertner, S. L., & Simons, R. F. (2011). Multiple group membership influences face-recognition: Recall and neurological evidence. *Journal of Experimental Social Psychology*, *47*(6), 1262–1268. doi:10.1016/j.jesp.2011.05.014

Herzmann, G., & Curran, T. (2013). Neural correlates of the in-group memory advantage on the encoding and recognition of faces. *PloS One*, *8*(12), e82797. doi:10.1371/journal.pone.0082797

Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2013). Same faces, different labels: generating the cross-race effect in face memory with social category information. *Memory & Cognition*, *41*(7), 1021–31. doi:10.3758/s13421-013-0316-7

Hugenberg, K., & Corneille, O. (2009). Holistic processing is tuned for in-group faces. *Cognitive Science*, *33*(6), 1173–1181. doi:10.1111/j.1551-6709.2009.01048.x

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychological Review*, *117*(4), 1168–1187. doi:10.1037/a0020463

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, *13*(4), 330–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5359231

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*(1), 3–35. doi:10.1037//1076-8971.7.1.3

Meissner, C. A., Brigham, J. C., & Butz, D. A. (2005). Memory for own- and other-race faces: a dual-process approach. *Applied Cognitive Psychology*, *19*(5), 545–567. doi:10.1002/acp.1097

Pauker, K., & Ambady, N. (2009). Multiracial faces: How categorization affects memory at the boundaries of race. *The Journal of Social Issues*, *65*(1), 69–86. doi:10.1111/j.1540-4560.2008.01588.x

Phillips, P. J., Beveridge, J. R., Draper, B. A., Givens, G., O'Toole, A. J., Bolme, D., Dunlop, J., Lui, Y. M., Sahibsada, H. & Weimer, S. (2012). The Good, the Bad, and the Ugly Face Challenge Problem. *Image and Vision Computing*, *30*(3), 177–185. doi:10.1016/j.imavis.2012.01.004

Pinter, B., & Greenwald, A. G. (2011). A comparison of minimal group induction procedures. *Group Processes & Intergroup Relations*, *14*(1), 81–98. doi:10.1177/1368430210375251

Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology*, *49*(2), 298–301. doi:10.1016/j.jesp.2012.10.017

Rhodes, G., Lie, H. C., Ewing, L., Evangelista, E., & Tanaka, J. W. (2010). Does perceived race affect discrimination and recognition of ambiguous-race faces? A test of the sociocognitive hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(1), 217–23. doi:10.1037/a0017680

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, *138*(1), 146–174. doi:10.1037/a0025750

Sangrigoli, S., Pallier, C., Argenti, A-M., Ventureyra, V. A. G., & de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science : A Journal of the American Psychological Society / APS*, *16*(6), 440–444.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law*, *7*(1), 36–97. doi:10.1037//1076-8971.7.1.36

Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: an event-related potential study. *Neuroreport*, *19*(5), 583–587.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178.

Van Bavel, J. J., & Cunningham, W. A. (2012). A social identity approach to person memory: group membership, collective identification, and social role shape attention and memory. *Personality & Social Psychology Bulletin*, *38*(12), 1566–78. doi:10.1177/0146167212455829

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: a functional magnetic resonance imaging investigation. *Psychological Science*, *19*(11), 1131–9. doi:10.1111/j.1467-9280.2008.02214.x

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience*, *23*(11), 3343–54. doi:10.1162/jocn_a_00016

Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological Psychology*, *89*(1), 137–147. doi:10.1016/j.biopsycho.2011.10.002

Williams, M. (2001). In whom we trust: Group membership as an affective context for trust development. *Academy of Management Review*, *26*(3), 377–396.

Young, S. G., Bernstein, M. J., & Hugenberg, K. (2010). When do own-group biases in face recognition occur? Encoding versus post encoding. *Social Cognition*, *28*(2), 240–250.

Young, S. G., & Hugenberg, K. (2010). Mere social categorization modulates identification of facial expressions of emotion. *Journal of Personality and Social Psychology*, *99*(6), 964–977. doi:10.1037/a0020400