RELATIVE DISTINCTIVENESS AND FACE RECOGNITION BIASES

By

Joyce M. Oates

Submitted to the

Faculty of the College of Arts and Sciences

of American University

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

In

Behavior, Cognition, and Neuroscience

Chair:

Z. Peynircioglu

Zehra F. Peynircioğlu, Ph.D.

Lynne Reder

Lynne M. Reder, Ph.D.

Scott Parker

Scott Parker

C. Stoodley

Catherine J. Stoodley, D.Phil.

Matthew G. Rhodes, Ph.D.

Dean of the College of Arts and Sciences

August 12, 2014

Date

2014

American University

Washington, D.C. 20016

This work is dedicated to Daniel Dickison.

RELATIVE DISTINCTIVENESS AND FACE RECOGNITION BIASES

BY

Joyce M. Oates

ABSTRACT

Face recognition biases occur when in-group (people with similar characteristics to the perceiver) faces are remembered better than out-group faces (Sporer, 2001). These biases have been observed for physical factors (e.g., race, age, gender) as well as psychological/social factors (e.g., occupation status; Ratcliff et al., 2011). Relative distinctiveness also affects face recognition such that atypical faces are better remembered than typical faces (e.g., Valentine, 1991). The Featural Fan Effect (e.g., Reder et al., 2002) can create stimuli that are relatively distinct without relying on atypicality, but rather by manipulating the number of items that share a given feature. In Experiment 1, we tested the effects of a psychological/social factor (occupation status) by labeling faces with a high- or low-status occupation, as well as the Featural Fan Effect by manipulating the number of times a given eye region appeared with faces. Curiously, there was no effect of occupation status, but a healthy effect of fan such that faces that did not share an eye region were better remembered than faces that did share an eye region. Since better memory for faces labeled with high-status occupations compared to low-status occupations has previously been reported (Ratcliff et al.), in Experiments 2a-2g we tested several psychological/social factors and compared them to two physical factors (age and gender) known to elicit face recognition biases. Although own-age and own-gender recognition biases occurred for the physical factors, we found no evidence to support differential memory based on group membership for psychological/social factors. Therefore, in Experiment 3, we once again tested the Featural Fan Effect by manipulating the number of times that an eye region was shared by faces, but this time used a physical factor of age (own-age: younger adult vs other-age: older adult). The results showed main effects of fan such that low-fan faces were better remembered than high-fan faces and of age such that own-age faces were remembered better than other-age faces. However, there was no interaction: the Featural Fan Effect did not moderate the effect of the own-age bias. The results are discussed within the frameworks of major models of face recognition biases.

ACKNOWLEDGMENTS

Many thanks to Daniel Dickison and Almas Kebekbayev for their invaluable help with computer programming.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF ILLUSTRATIONS

Figure

CHAPTER 1

RELATIVE DISTINCTIVENESS AND FACE RECOGNITION BIASES

Humans are generally experts at recognizing human faces. There are several factors that influence such recognition, exemplified by a variety of face recognition biases. For instance, the own-race bias (e.g., Meissner & Brigham, 2001), the own-gender bias (e.g., Cross, Cross, & Daly, 1971; Lovén, Herlitz, & Rehnman, 2011; Wolff, Kemter, Schweinberger, & Wiese, 2013), and the own-age bias (e.g., Harrison & Hole, 2009; M. G. Rhodes & Anastasi, 2012) are characterized by better performance, that is they have greater memory sensitivity for in-group versus out-group members on recognition tasks. In-group members share similar characteristics with the perceiver, such as race, age, gender, political affiliation, whereas out-group members do not.

Perhaps the best-known face recognition bias is the own-race bias. Typically, face recognition data show a mirror effect (which says that in recognition memory, performance on previously unseen items "mirrors" performance on previously seen items; Glanzer and Adams, 1985) and are characterized by lower hit rates and higher false alarm rates for other-race faces (Bruce & Young, 2012). For example, Caucasians are better at recognizing Caucasian faces than Asian faces, and vice versa (Michel, Rossion, Han, Chung, & Caldara, 2006). One overarching explanation offered for the own-race bias (as well as face recognition biases in general), which underlies the more specific theories to be discussed later, assumes that own-race faces are processed holistically whereas other-race faces are processed piecemeal or feature by feature (e.g., Meissner & Brigham, 2000). Holistic processing refers to perceiving the face as a Gestalt, a unified image, and piecemeal processing refers to perceiving the face as the compilation of a set of features. Evidence for holistic processing comes from inversion and composite tasks (e.g., Michel et al., 2006; Young, Hellawell, & Hay, 1988), in which participants are faster and more accurate at recognizing upright compared to inverted faces and are better at identifying the top or bottom half of composite faces when they are misaligned (when the top and bottom half of a face composite are shifted so that the face is no longer seen as a continuous whole) compared to when they are intact. Thus, because own-race faces are processed holistically where the whole is assumed to be greater than the sum of the features[1], inverting or misaligning them disrupts this Gestalt (G. Rhodes, Hayward, & Winkler, 2006). In composite tasks, the face Gestalt actually works *against* identifying halves of a face that are aligned (and thus form a whole face image) by superseding feature by feature processing compared to those faces that are misaligned (Young et al., 1988). When the task is to determine whether the top or the bottom half of the face is from the target that was just studied as a complete face, performance improves (participants become faster, are less error prone) more for shifted own-race than shifted other-race faces (Michel et al., 2006) suggesting that aligned own-race faces are more likely to be processed as a Gestalt than as a set of features.

[1] This claim has been contested by Gold, Mundy, and Tjan (2012) who used a Bayesian Optimal Integrator to model their data and demonstrated that a whole face is not processed in a manner that is greater than the sum of processing each feature individually.

As in the own-race bias, inversion effects have also been found in the own-age bias, the bias that is examined in the current study. Whereas young adults show an inversion effect for own-age faces, they do not for infant faces, and they exhibit a diminished inversion effect for child faces (Kuefner, Macchi Cassia, Picozzi, & Bricolo, 2008). That is, the inversion effect arises even though the own-age bias might be grounded in a different mechanism than that for the own-race bias. Indeed, whereas most face recognition biases remain somewhat stable across time (e.g., a young adult who performs poorly with other-race faces now will probably perform poorly with other-race faces years from now), the own-age bias does not because the in-group changes as people age (Hills, 2012). In a longitudinal study, Hills (2012) found that children at age 8 were more likely to recognize faces of their own age compared to faces of children age 7 or 9 years old. Hills suggests that people rapidly adapt to the own-age bias as they themselves grow older; an 8-year-old child who is an expert at recognizing pictures of 8-year-old faces does not maintain this expertise one year later. Presumably, for adults, this window of adaptation in recognizing other adults is not as narrow as it is in children and adolescents, since physical changes in adulthood are not as dramatic and occur over longer spans of time. Further, in their meta-analysis, M. G. Rhodes and Anastasi (2012) note that it is the recent experience, not lifetime experience, that is more likely to drive the own-age bias. Regardless of whether or not the mechanism for the own-age bias differs from that for the own-race bias, however, the patterns of results in face recognition studies and tests of face inversion are similar in the two biases.

There are several theories that attempt to explain usually one (the own-race bias) but sometimes more (including also the own-age bias) of these recognition biases. The current major theories[2] include the Experience-Based Holistic Theory (Rossion & Michel, 2011), the Face Space Model (Valentine, 1991), the In-group/Out-group Theory (Sporer, 2001), and the Categorization-Individuation Model (Hugenberg, Young, Bernstein, & Sacco, 2010).

The Experience-Based Holistic (EBH) Theory proposed to explain the own-race bias (Rossion & Michel, 2011) states that as humans develop expertise in face processing, their perceptions for faces become finely tuned for a specific type of face, namely the type they are exposed to the most: own-race faces. Therefore, our processing difficulties with other-race faces are due to our lack of perceptual experience with them and the finely tuned own-race face perception system. We have one holistic face template and only faces that match that template are processed holistically. According to EBH, although it is possible to change the ability to process other-race faces based on situations and socially motivated face learning, the extent to which face recognition is altered by those factors is small. That is, EBH does not account for data that show that holistic/Gestalt face processing can be modified quite a bit by psychological/social factors (e.g., faces that are experimentally assigned as belonging to more powerful or more important people). For example, it cannot account for data

[2] There is one other major theory in the literature, the own-race bias contact hypothesis, which is excluded here because, although included as a component in face recognition biases (Hugenberg et al. 2010; He, Ebner, & Johnson, 2011; Wiese, 2012), it does not appear to be complete, in and of itself, in explaining the complexity of face recognition biases. Further, in some instances, it is questionable whether or not it may be a component in all face recognition biases, such as the own-gender bias (Wright & Sladden, 2003).

where own-race and own-age faces are affected by in-group/out-group membership in a social community (Hugenberg & Corneille, 2009). Thus, although EBH is a useful model for own-race template development, it does not explain recognition biases that do not originate from perceptual experience.

Valentine's (1991) Face Space Model assumes that faces are coded in space on a graphical representation of unspecified n-dimensions. Faces that resemble the norm or an exemplar face (a similar face template idea is also a driving factor in EBH) are widely dispersed around the origin of the n-dimensional space. Since this particular set is widely spread, these faces dispersed around the origin are distinctive, and thus it is relatively easy to discriminate an individual face. This does not hold true for faces that do not resemble the norm or exemplar faces. Their graphical depiction in the n-dimensional space is located within a restricted range at a distance from the origin. Because expertise with those non-exemplar faces does not allow for a wide range of discrimination based on the feature dimensions, the faces tend to be are grouped closely together. This grouping makes it difficult to discriminate any individual face and thus the Face Space Model can account for the own-race bias and inversion effects. However, a major criticism put forth by Valentine himself is that the model is underspecified. The type and number of dimensions in the face space and the scales of measure for those dimensions are undefined. It is not known what kind and how many dimensions are needed to comprise an exemplar face.

Given this issue, and given that, as with EBH, there is no provision for psychological/social group factors of face recognition, it is not yet an optimal model for testing explicit parameters of generic face recognition biases.

The basic premise of the In-group/Out-group Theory (Sporer, 2001) is that perception of visual cues in faces (e.g., color of skin) determines whether or not a face will be distinguished beyond the category level (e.g., Caucasian face). If a face is determined to be a member of the in-group, then it is processed more deeply and in a holistic manner. If a face is determined to be a member of an out-group, it is processed more shallowly and in a piecemeal manner. Therefore, in-group faces are afforded a memorial advantage. However, this theory does not include face perception development which makes the model unable to account for face recognition biases that are not based on contact alone, such as the own-race bias in 6-9 month old infants (e.g., Kelly, Quinn, Slater, Lee, Ge, & Pascalis, 2007) and evidence that exposure to other-race faces does not decrease the own-race bias (e.g., Rossion and Michel, 2011).

Thus, all of these theories, except the Categorization-Individuation Model (CIM; Hugenberg et al., 2010), appear to be underspecified, and may not be suitable as an explanation for generic face recognition biases. CIM entails three separate but somewhat related factors that influence face recognition: perceptual experience, social categorization, and motivated individuation. First, people typically develop experience with a single race that enables perceptual abilities to become more finely tuned for that race. Although this is not necessarily the case for own-age faces because our age is constantly changing, perceptual experience nevertheless can contribute to the own-age bias as well (M. G. Rhodes & Anastasi, 2012). In addition, expertise with other-age faces can be developed, as is often the case with elementary school teachers (Harrison & Hole, 2009) and maternity-ward nurses (Macchi Cassia, Picozzi, Kuefner, & Casati, 2009) who show evidence of holistic face processing for other-age children/baby faces.

Second, people can be categorized as “own” or “other” based on physical factors such as gender, age, and race. In cases where the person’s face belongs to an out-group category, physical categorization (e.g., labeling the race of the face) without individuation is the default. The categorization process is automatic (Caldara, Rossion, Bovet, & Hauert, 2004; Levin, 1996; 2000) and effortless whereas individuation is effortful, and people tend to individuate faces only of their in-group because they have more expertise and more motivation to do so because of an “us” vs “them” frame of mind. Since out-group faces are categorized at a superordinate level (cf. Mervis & Rosch, 1981), it is more difficult to discriminate between other out-group faces. For instance, labeling each face in a set of faces as “Asian” would not help distinguish them from other faces within the same category.

Finally, motivated individuation drives whether or not an out-group face is remembered later on. For instance, despite similar levels of other-race contact for both Caucasian and African American participants, Caucasians show a greater own-race bias than African Americans (Meissner & Brigham, 2001). Hugenberg et al. (2010) propose that this may be because Caucasians typically have higher socio-economic status (SES) than African Americans. As a result, participants are more motivated to pay attention to people with greater SES and therefore have greater incentive to remember Caucasian faces (Hugenberg et al.). Alternatively, where most of this research has been done, Caucasians have typically comprised the majority, and minority groups may be more motivated to pay attention to faces in the majority. For example, historically, most Hollywood movies and American/British (English speaking) television shows have typically featured only Caucasians in lead roles.

According to the CIM, motivation to individuate can apply to different sub-groups of in-group faces, as well (e.g., Hugenberg & Corneille, 2009). With own-race faces, Hugenberg and Corneille showed larger composite effects when faces were categorized as from the same university (in-group) versus from a different university (out-group). CIM also predicts that people of greater SES should be better remembered because there is greater motivation to individuate them and indeed, Ratcliff, Hugenberg, Shriver, and Bernstein (2011) found better face recognition for face stimuli labeled with higher status occupations compared to lower status occupations. The literature to date shows that CIM is the only model that accounts for recognition biases not reliant on physical factors (such as age, race, gender).

To summarize, despite minor differences in all these face recognition biases, the commonality is that there is better recognition of in-group than out-group faces. This commonality is typically explained by the idea that in-group faces are processed as Gestalts, whereas out-group faces are not processed as efficiently as one whole, but rather parsed feature by feature (e.g., Hugenberg & Corneille, 2009; Rhodes, Brake, Taylor, & Tan, 1989; Tanaka, Kiefer, & Bukach, 2004). Within this context, we focus on one of the major face recognition bias models (Categorization-Individuation Model;

Hugenberg et al., 2010), which appears to be the most complete and testable of the models. One factor that this model does not incorporate among its explanations of face recognition biases, however, is the idea of relative distinctiveness of the faces themselves and the role such distinctiveness plays in differentiating in-group and out-group faces. Distinctiveness affects face recognition in general (e.g., Going & Read, 1974; Leder & Bruce, 1998; Vokey & Read, 1992) and the prediction is that it would also affect face recognition biases.

When stimuli are relatively distinct from their neighbors or other competing stimuli, they are often perceived and remembered better (for a review see Hunt, 2013). Several findings and theories such as the Cue Overload Principle (e.g., Watkins & Watkins, 1975), the Von Restorff Effect (e.g., Huang & Wille, 1979; Von Restorff, 1933), SIMPLE (Scale Independent Memory and Perceptual Learning; Neath & Brown, 2006), and the Featural Fan Effect (e.g., Reder, Donavos, & Erickson, 2002) successfully address this idea. In regards to faces, the general finding is that perceptual and memorial performance is better for faces that are distinctive/atypical (e.g., Going & Read, 1974; Leder & Bruce, 1998; Vokey & Read, 1992). Within this realm, distinctiveness has been defined either as an extra-experimental factor (based on participants' experience)—for instance, faces rated atypical were found to be better remembered, (e.g., Vokey & Read, 1992)—or as an exaggeration of features as a caricature, which were better remembered than un-exaggerated faces (e.g., Bruce & Young, 2012; Leder & Bruce).

In this study, the primary purpose was to determine if relative distinctiveness would indeed moderate face recognition biases. We use the Featural Fan Effect (e.g., Anderson & Paulson, 1978; Reder et al., 2002) to manipulate distinctiveness. Faces are not identified or recognized in isolation, but rather against a backdrop of other competing faces. Further, faces are often identified within a homogeneous group (e.g., all Caucasian females). Thus, an important question is what enables some faces rather than others to be successfully identified and/or remembered. Also, we test this idea within the CIM framework because it is the most comprehensive of the models to date.

The Featural Fan Effect occurs when memory for some items varies as a function of how many other items are associated with their respective contexts. It is dependent upon the reinstatement of the encoding context as reflected by the encoding specificity principle (Tulving & Thomson, 1973). For instance, Graf and Ryan (1990) showed that, under certain conditions, words tested in their reinstated font were better remembered than if they were tested in a different font. Further, the memorial benefit of a reinstated context (perceptual features) is influenced by how many items are associated with that context (Diana, Peterson, & Reder, 2004a; Reder et al., 2002; Park, Arndt, & Reder, 2006). The general finding is that test items reinstated (presented in their original context) in a low-fan feature (feature shared with few other items) are remembered better than items reinstated in a high-fan feature (feature shared with many other items). For example, words presented in unique fonts are remembered better than words presented in a font that had been presented with many other words (Diana et al., 2004a; Reder et al., 2002; Park et al., 2006). In this case, the context is the font and the number of words the font appears with is the fan. The Featural Fan Effect has typically been demonstrated with words using perceptual dimensions such as font, color, and voice (e.g., Park et al., 2006) and Reder et al. (2013) has also demonstrated this effect with well-known backgrounds and famous people. Reder et al. (2013) showed that the number of famous people associated with a given background can affect memory for those people later on.

In the only experiment that manipulated the features of faces themselves to vary the distinctiveness of the faces, Anderson and Paulson (1978) found an effect of fan on recognition response times for features in composite faces. Their materials were IDENTI-KIT sketches comprised of four features: hairstyle, eyes and eyebrow region, nose and mouth region, and the chin region. Participants studied faces that had two high-fan features (features seen with three other faces), two low-fan features (features seen with only one face each), and faces that had one high feature and one low-fan feature. The results showed that participants were slowest to recognize faces with two high-fan features, which supported the Featural Fan Effect. Anderson and Paulson also attributed the weaker effect of fan with faces compared to verbal information to the stronger Gestalts formed by the faces. That is, it was assumed that when it was not explicitly pointed out to the participants which features were being manipulated (e.g., only the hairstyle; Anderson & Paulson, Experiment 2), participants may have defaulted to holistic processing rather than piecemeal processing.

If this attribution of stronger Gestalts with faces is correct and if out-group faces are processed piecemeal, then these out-group faces should show a greater effect of fan manipulation than in-group faces. However, if fan moderates the number of similar-looking faces, and in-group faces are more likely to be processed as Gestalts, then in-group faces should show a greater effect of fan than out-group faces. Whereas in-group faces might be classified/individuated by fan of similar-looking faces (high-fan = many similar faces, low-fan = unique face), out-group faces might be lumped into a single category (out-group face—they all look the same), and not further individuated. Therefore, since faces would not be individuated, performance on out-group faces should be the worst overall but should show less of an effect of fan than in-group faces. Consistent with this prediction are studies on the effects of context and face recognition biases (e.g., Horry & Wright, 2008; Horry, Wright, & Tredoux, 2010), which have shown that background pictures are better remembered when presented against in-group than out-group faces. That is, faces that show less of a context effect might also show less of a Featural Fan Effect. Finally, the last possible outcome is that if all faces, regardless of group membership, are strongly processed as Gestalts, an effect of fan may not emerge. There is a possibility that pictures of real faces may be formed as stronger Gestalts than the composite sketches used by Anderson and Paulson (1978).

Anderson and Paulson (1978) showed a Featural Fan Effect with sketches of faces, and Reder et al. (2013) showed a Featural Fan Effect with faces when the fan of the backgrounds was varied (i.e., better memory for faces presented on low-fan backgrounds). However, the latter finding was found only when the faces belonged to famous people, a result that supported Reder et al.'s assertion that pre-existing memory traces were easier to associate with a context. On the other hand, Reder et al. (2013)

did not find a Featural Fan Effect with non-famous faces and, further, in a study more pertinent to present purposes, Diana and Reder (2004b) did not find a typical fan effect when they tested for it within the context of the own-race bias. In that study, fan of clothing (hat, scarf, jacket, etc.) was crossed with race (Caucasian-own or African American-other).Although the faces presented with reinstated clothing were indeed remembered better, there was no evidence of the Featural Fan Effect for own-race faces. That is, own-race faces presented with low-fan clothing were not remembered better than own-race faces presented with high-fan clothing. This was surprising because in previous work with verbal materials, Reder and colleagues had found that low-fan fonts had benefitted word recognition (e.g., Diana et al., 2004a; Reder et al., 2002; Park et al., 2006).

Although Reder and colleagues did not find any evidence for the own-race bias while testing for fan effects (Diana & Reder, 2004b), nor a fan effect for unfamiliar faces (Reder et al., 2013), there were several reasons why these effects may not have emerged. In both cases, the study stimuli were presented for only 2 s, whereas other research using unfamiliar faces had employed longer study times of 5-6 s and fewer stimuli (e.g., Tanaka & Farah, 1993; Toseeb, Keeble, & Bryant, 2012). Indeed, not only did participants have to encode each face within 2 s, but also had to encode an associated context during that time as well. In addition, since membership of in-group/out-group was not a question of interest in the Reder et al. (2013) study, participants were of mixed race and gender and stimuli were of mixed age, race, and gender. Therefore, at least for the unfamiliar faces, it is possible that poorer performance with other-group stimuli could have diluted the memory benefit for the own-group stimuli. Thus, in the present study, participants were kept more homogeneous with respect to the in-group faces, that is, we used stimuli that were the same age as the participants. In addition, participants were given longer study times and more opportunities to learn the faces.

In the current study, we examine how relative distinctiveness, operationalized as featural fan, might moderate recognition biases. In Experiment 1, the recognition bias was the psychological/social factor of occupation status (high vs low) in accordance with the specifications of CIM (Ratcliff et al., 2011). For the Featural Fan Effect, we manipulated the number of times that a given eye region was presented with different faces. The eye region was chosen as the internal feature to manipulate because it appears to be the most important for face identification (e.g., Eimer, 1998; Johnson, 2005; Letourneau & Mitchell, 2008; Valentine & Bruce, 1985; Whalen et al., 2004). Eimer (1998) demonstrated that an N170 (an event-related potential regarded as a neural marker for face processing) emerged for eye regions in isolation (without the face), but that faces presented without eye regions showed a delay in that component. Letourneau and Mitchell (2008) observed faster latencies for the N170 for the top halves (the eye region) than the bottom halves of composite faces, suggesting that eyes are processed automatically. Also, within the context of a face, the eyes appear to drive the Thatcher Effect (Valentine & Bruce, 1985). The Thatcher Effect is a phenomenon that occurs when the internal features of a face are inverted. The inversion is less likely to be noticed (less likely to appear abnormal) when the head is displayed upside down compared to when the head is displayed upright (Thompson, 1980). That the eyes alone can drive the effect indicates that they may be the most important internal face feature. Lastly, newborns tend to orient more towards features on the top half of the face (e.g., Johnson, 2005) indicating a preference for looking at eyes. Therefore, it appears that humans are specialized for processing the eye region.

In addition to testing a psychological/social group factor, with respect to the Featural Fan Effect, we extend the previous work of Anderson and Paulson (1978) to photographs of real faces. The Fan Effect would predict that faces whose eyes were a low-fan feature should be better remembered than faces whose eyes were high-fan. CIM (and general social-cognitive accounts of face recognition biases) would predict that the group manipulation should have an impact such that 1) faces labeled with high status-occupations would be individuated and therefore remembered better, and 2) faces labeled with low-status occupations would not be individuated and therefore remembered worse (cf. Ratcliff et al., 2011). Further, since in-group (high-status occupations) faces are individuated, this would suggest that a Featural Fan Effect would emerge for those stimuli whereas a Featural Fan Effect would not emerge for out-group (low-status occupations) faces, which are not individuated.

Interestingly, the results of Experiment 1 were, in fact, contrary to CIM's prediction and we did not observe an effect of occupation status. Therefore, we further examined psychological/social group factors as well as physical group factors in Experiments 2a-2g. Because we did not find an effect of group for any of the psychological/social factors in Experiments 2a-2e, in Experiment 3 we once again examined the Featural Fan Effect of the eye region but this time used a physical group factor of age (own-age: younger adults vs other-age: older adults) to look at whether distinctiveness would moderate the Featural Fan Effect.

CHAPTER 2

EXPERIMENT 1

The aim of Experiment 1 was to test 1) the Featural Fan Effect by manipulating the number of times eye regions (eyes and brow) are presented with faces and 2) the psychological/social group factor of occupation status (high-status vs low-status). In order to control for extra-experimental (personal experience) effects of out-group, the Caucasian face stimuli comprised the same age and gender as the participants. It should be noted that recognition biases do not appear to affect faces that have preexisting representations (e.g., Phelps et al., 2000). That is, familiar faces that are already part of our memory system due to multiple previous exposures, such as those of our close friends or celebrities, appear to be immune to recognition biases. Therefore, all faces in the present experiments were unfamiliar to the participants. However, because recognition of unfamiliar faces presented once for only 2 s has not shown a Featural Fan Effect (Reder et al., 2013), the familiarization process during the experiment was increased (cf. Bonner, Burton, & Bruce, 2003) such that participants were exposed to each face on three separate occasions for a total of a minimum of 7 seconds.

METHOD

Participants. Sixty-five female participants from American University between ages 18-22 (mean age=19.4) received 1 extra credit for their Psychology courses. Participants were screened for age and gender but not race in order to determine if an ad-hoc grouping by same- or other-race might reveal any moderating effects. As revealed in an exit survey, 37 were of the same race as the people depicted in the stimulus pictures whereas 28 were of a different race.

Figure 1. Example of How Eye Regions were Masked to Create Composite Faces. From left to right at the top: original face 1, original face 1 with the eye region masked (such that it is transparent). From left to right at the bottom: original face 2, original face 1 with the eye region from original face 2. All stimuli were edited such that no single original face appeared intact. In addition, each participant received a different random order of stimuli pairings. Any oddities due to image manipulation that made certain combinations more distinctive added to the noise/error term.

Materials. The face stimuli were generated from 112 pictures of Caucasian females between ages 17-31 (mean age=20.9) were obtained from the Glasgow Unfamiliar Face Database (Burton, White, & McNeill, 2010) as well as Flickr.com.

Images obtained from Flickr.com were Creative Commons-licensed content that could be modified, adapted, or built upon. All images were photo-edited to create composites using Pixelmator. More specifically, the eye regions were masked from the original faces and then combined with the original faces to create the composite faces. Thus, there were a total of 12,432 (111 x 112) composite face stimuli, and no faces remained in their original form in the composite set. Examples of two original faces, and one composite face in which the rest of a face was combined with a new eye region, are shown in Figure 1. Superlab 5.0 was used to code the experiment script.

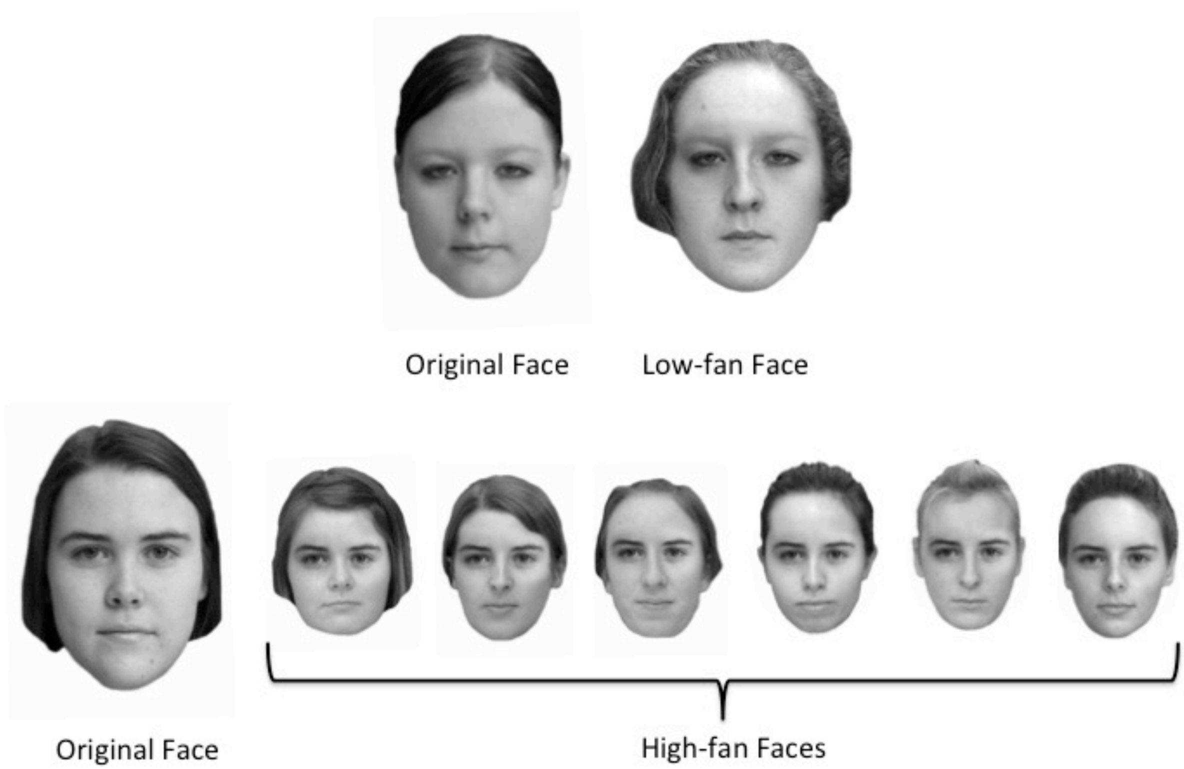


Figure 2. Examples of Low- and High-fan Faces. The high-fan faces all shared the same eye region.

Design & Procedure. Experiment 1 consisted of a 2 (fan type: low vs high) x 2 (occupation status: high vs low) repeated measures design. A given eye region was presented with just one face in the low-fan condition or with six faces in the high-fan condition (see Figure 2 for an example of the stimuli). Stimuli compilation (combinations of faces and eye regions) and presentation order were randomized with respect to both fan type and occupation status conditions. For each participant, any given face had an equal opportunity to be in the low or high fan condition as well as in the high or low status occupation condition. During the study phase, overall, each participant saw 112 pictures (see Figure 3 for a graphical depiction of the study phase design).

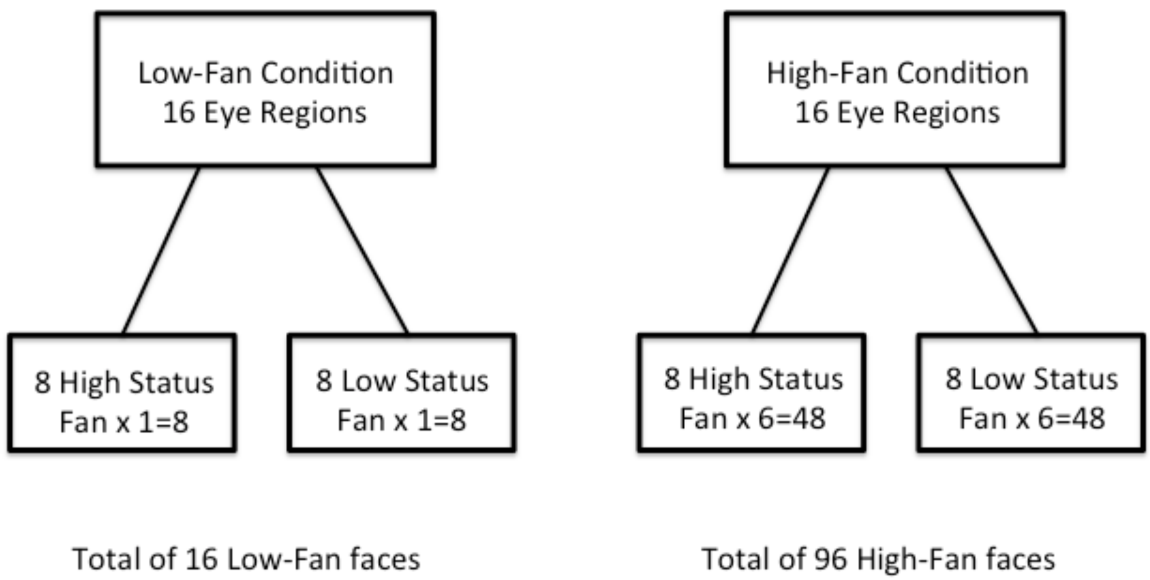


Figure 3. Design of the Study Phase for Experiment 1.

For the manipulation of fan, there were 16 eye regions assigned to the low-fan condition and 16 eye regions assigned to the high-fan condition. Thus, the low-fan condition consisted of 16 different eye regions composited with 16 different faces. The high-fan condition consisted of 16 different eye regions, each with a fan of 6 (each eye region was composited with 6 different faces) for a total of 96 different faces. For the manipulation of occupation status, half of the low-fan and half of the high-fan faces were assigned to the high-status condition (doctor, CEO, lawyer, professor) and the rest were assigned to low-status condition (maid, nanny, secretary, bus driver).

The occupation labels always appeared below the faces, although no special orienting task was implemented to draw special attention to them so as to replicate the methodology of Ratcliff et al.’s (2011) study. Some of the occupations were taken from Ratcliff et al., (high-status: doctor, CEO, four-star general; low status: mechanic, fry cook, farmer, plumber) and others were created new for this experiment. In order to ensure that high-status occupations were indeed believed to be higher status than low-status occupations, ratings were collected during an exit survey where on a scale of 1 to 7, 1 indicated the lowest status and 7 indicated the highest status (cf. Ratcliff et al.).

The study phase itself was further divided into a training task and a matching task. Once again, in accordance with the method by Ratcliff et al. (2011), there were no explicit instructions in the study phase to encode the occupation labels with the faces. During the training task, participants studied the faces that were presented randomly, one at a time on a computer screen for 5 s each, with an inter-stimulus interval (ISI) of 250 ms during which a fixation cross “+” was presented. Then, participants completed the matching task with pairs of all the faces that were presented in the training task. The purpose of the matching task was merely to provide more opportunity for participants to become familiar with the target faces. During this task, the two faces of a given pair were presented side by side, and participants determined, as quickly and as accurately as possible, whether or not the two faces were the same or different. They responded with the “s” key if the two faces were the same or the “d” key if the two faces were different. Participants were instructed to make button presses with their dominant hand. In order to facilitate learning, the computer program presented immediate feedback regarding accuracy. If the participant answered correctly, “Correct” in green appeared onscreen. If the participant answered incorrectly, “Incorrect” in red appeared onscreen. Initial observations showed that incorrect answers were quite rare. The stimuli were randomly paired for each participant within the constraint that a given face was paired with itself and also paired with a different face. In this way, participants had three opportunities to learn the faces, once in the training task and twice in the matching task.

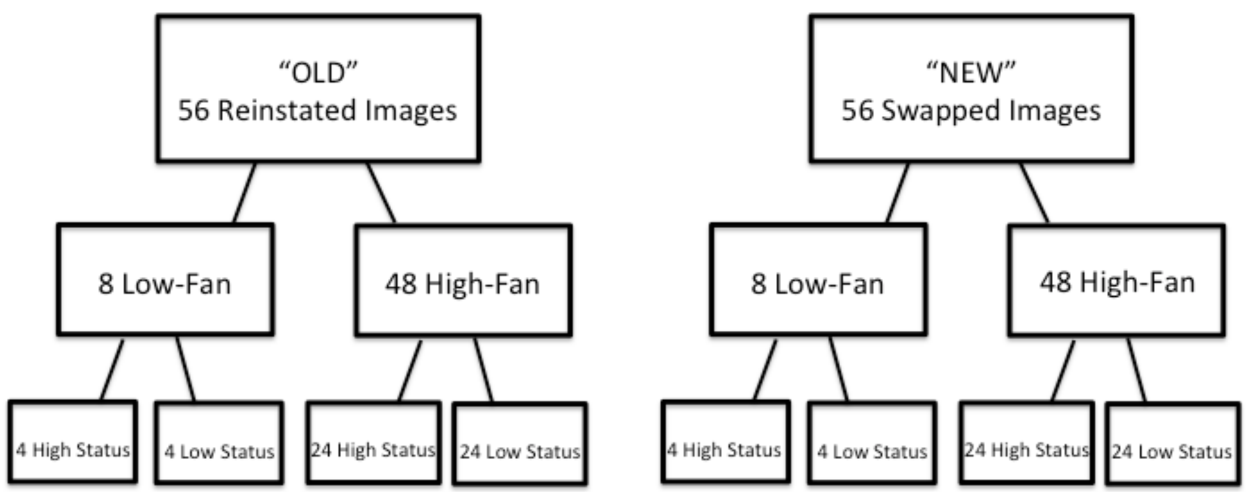


Figure 4. Design of the Test Phase for Experiment 1.

Following the study phase directly, a free-choice recognition was given (see Figure 4 for a graphical depiction of the test phase design). Half of the faces were reinstated (presented exactly as they were presented at study), and called “Old”. The other half consisted of faces with swapped/re-paired, eye regions. These were called "New" because the exact face-eye region pairing had not been seen before. Reinstated and swapped faces were randomly determined for each participant within the constraint that there were equal numbers of each type. Further, swapped faces were randomly re-paired for each participant with the constraint that eye regions were swapped within the study condition. That is, a face previously presented with a low-fan eye region was presented with a different low-fan eye region and 6 faces presented with a high-fan eye region were presented with a different high-fan eye region. Participants were instructed to discriminate whether an exact face had been presented in the previous tasks based on recognition of the whole face. Therefore, the test phase also consisted of 112 stimuli (along with their occupation labels): 56 reinstated or "Old" faces (8 in the low-fan and 48 in the high-fan conditions), and 56 swapped or "New" faces (8 in the low-fan and 48 in the high-fan conditions). As in the study phase, these test stimuli were presented in random order with an ISI of 250 ms during which a fixation cross "+" was presented. Responding was self-paced, although participants were told to respond quickly without losing accuracy, using their dominant hand. Each test face remained onscreen until the participant made a judgment by pressing either the "j" key which was covered with a sticker labeled "Old" or the "k" key, which was covered with a sticker labeled "New". Measures of interest for this experiment and all subsequent experiments included hits and false alarms for the recognition test. Hits and false alarms are reported in Appendices A and B.

RESULTS

Manipulation check for rank of occupations. First, a paired-samples t-test revealed that, as expected, high-status occupations (mean=6.4), were rated significantly higher than low-status occupations (mean=3.0), $t(64)= 35.6$, $p<.001$, Cohen's d=4.6. Signal detection analyses were then conducted as in all following experiments, as well, yielding scores for d′ prime and C[3].

D′ results. The data are presented in Figure 5.

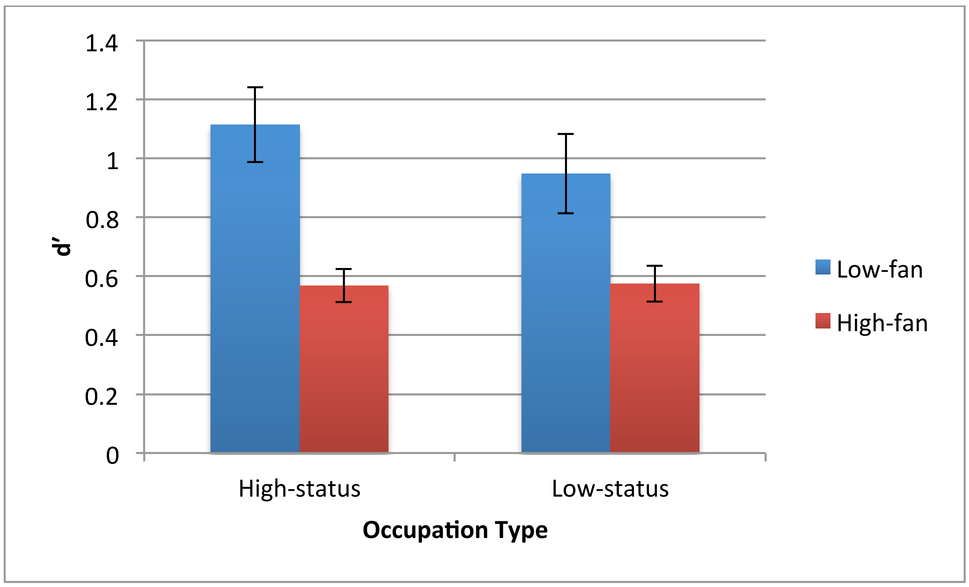


Figure 5. Experiment 1, Mean d′ as a Function of Occupation and Fan Type. Error bars represent the standard error of the means.

A fan type (low vs high) x occupation status (high vs low) x race (same vs other) repeated measures ANOVA revealed no effect of race, $F<1$, and no interactions between race and fan type or occupation status, $F(1, 63)=1.6$, $p>.10$, and $F<1$, respectively. Thus, we removed race from the model, and a fan type (low vs high) x

[3] The criterion C was calculated as follows: C = -[z score(hits) + z score(false alarms)]/2

occupation status (high vs low) repeated measures ANOVA was conducted. There was a main effect of fan type, $F(1, 64)=25.8$, $p<.001$, partial $\eta^2=.29$ such that low-fan faces were remembered better than high-fan faces, but no effect of occupation status or an interaction between fan type and occupation status (both *F*s<1).

C results. The data are presented in Figure 6. Similar results were obtained with the omnibus ANOVA and we removed race from the model to conduct a fan type (low vs high) x occupation status (high vs low) repeated measures ANOVA. There were main effects of fan, $F(1, 64)=61.7$, $p<.001$, partial $\eta^2=.49$ and occupation status, $F(1, 64)=3.9$, $p<.05$, partial $\eta^2=.06$. There was also a marginal interaction between fan type and occupation status, $F(1, 64)=3.1$, $p=.08$, partial $\eta^2=.05$.

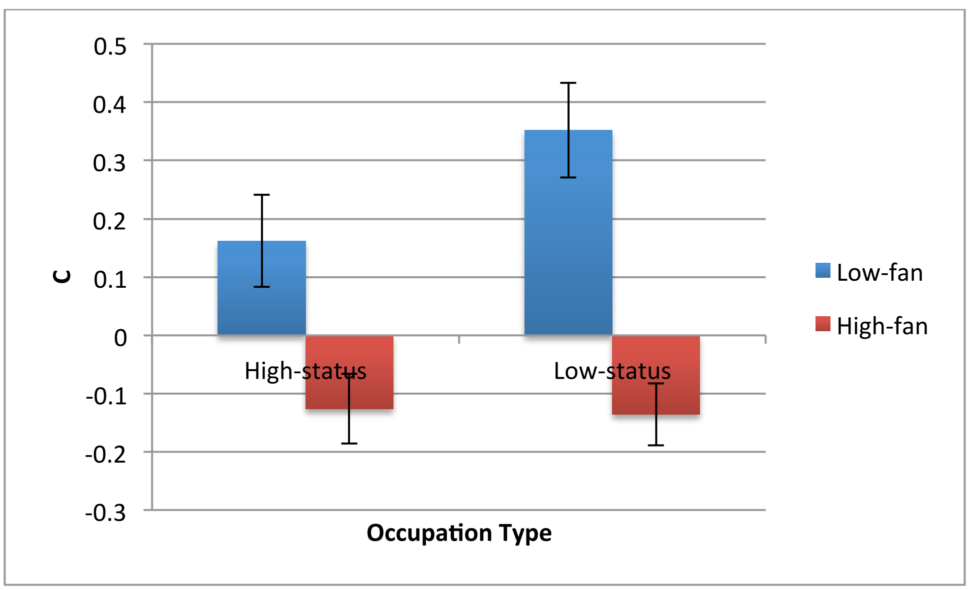


Figure 6. Experiment 1, Mean C as a Function of Occupation and Fan Type. Error bars represent the standard error of the means.

Thus, d′ scores revealed no evidence of better memory performance for high-status compared with that for low-status occupations. However, a strong fan effect emerged such that faces with a low-fan eye region were remembered better than faces with a high-fan eye region. C scores revealed an effect of fan such that participants were more likely to respond “Yes” to indicate that they had previously seen that exact person to high-fan faces than low-fan faces; that is they were more likely to respond liberally. Participants were also more likely to respond “Yes” to high-status occupations than low status occupations. These findings are inconsistent with CIM and the results that support CIM, reported by Ratcliff et al. (2011), where a larger d′ was observed for high-status occupations and there was no difference in C between high- and low-status occupations. To the contrary, we did not find a larger d′ for high-status occupations. With respect to the Featural Fan Effect, because CIM has no provisions for relative distinctiveness, it is not obvious how the model would predict better memory for low-fan faces and more liberal responding to high-fan faces. Perhaps the model could explain more liberal responding to out-group faces as evidence for lack of motivation to individuate those faces. However, Ratcliff et al. reported no such criterion shift for low-status (out-group) occupations and in fact, we found a trend in the opposite direction.

CHAPTER 3

EXPERIMENTS 2A-2G

Curiously, we did not find an effect of better memory performance for high-status occupations in Experiment 1. Therefore, the purpose of this next set of experiments was to test the predictions of CIM that both psychological/social and physical factors can produce equally robust recognition biases. We did not use a fan manipulation in Experiments 2a-2g, and thus only pictures of original faces were presented.

In each of the following experiments, half of the faces were randomly assigned for each participant to belong to the in-group condition and the rest to the out-group condition to determine whether an in-group recognition bias would arise as a function of the particular factor being manipulated.

Experiment 2a

Experiment 2a tested the effect of psychological age (labeling a face as 20 years old vs 30 years old, a novel factor that had not been previously tested).

METHOD

Participants. Thirty Caucasian female participants from Amazon's Mechanical Turk (AMT) between ages 20-22 (mean age=21.3) received a payment of $2. In this experiment and other experiments using AMT, participation was restricted to Caucasian females living in the United States. None of the participants had participated in the previous experiment.

Materials. The stimuli consisted of 100 pictures of young (mean=22 years), Caucasian, female faces collected from the Glasgow Unfamiliar Face Database (Burton et al., 2010).

Design. The factor of Group (psychological age) was manipulated such that half of the stimuli were labeled with “20 years old” (in-group) and the rest as “30 years old” (out-group). The in-group and out-group faces were assigned to either a blue or grey colored background in order to give an additional indicator of group membership (Ratcliff et al. 2011 used background color in Experiment 1b to indicate group membership). Group assignment to background was randomized for each participant. The background color always corresponded to the group label, that is, if “20 years old” was assigned to blue backgrounds, faces with that label always appeared on blue backgrounds (in both the study and test phases). There were 25 in-group and 25 out-group faces presented at study. At test, all previously seen faces were presented along with an equal number of unseen in-group and out-group faces for a total of 100 test stimuli. In order to eliminate confounds from item combinations and order effects, faces, labels, and background colors were randomly assigned for each participant.

Procedure. Participants were instructed that they would first complete a study phase followed by a subsequent recognition test for their memory of the faces presented in the study phase (see Figure 7 for examples of stimuli used in all these experiments and procedure). During the study phase, participants were instructed to rate the faces for intelligence (“how intelligent do you think the person you just saw is?”) using a scale of 1 to 4 (1=very unintelligent, 2=somewhat unintelligent, 3=somewhat

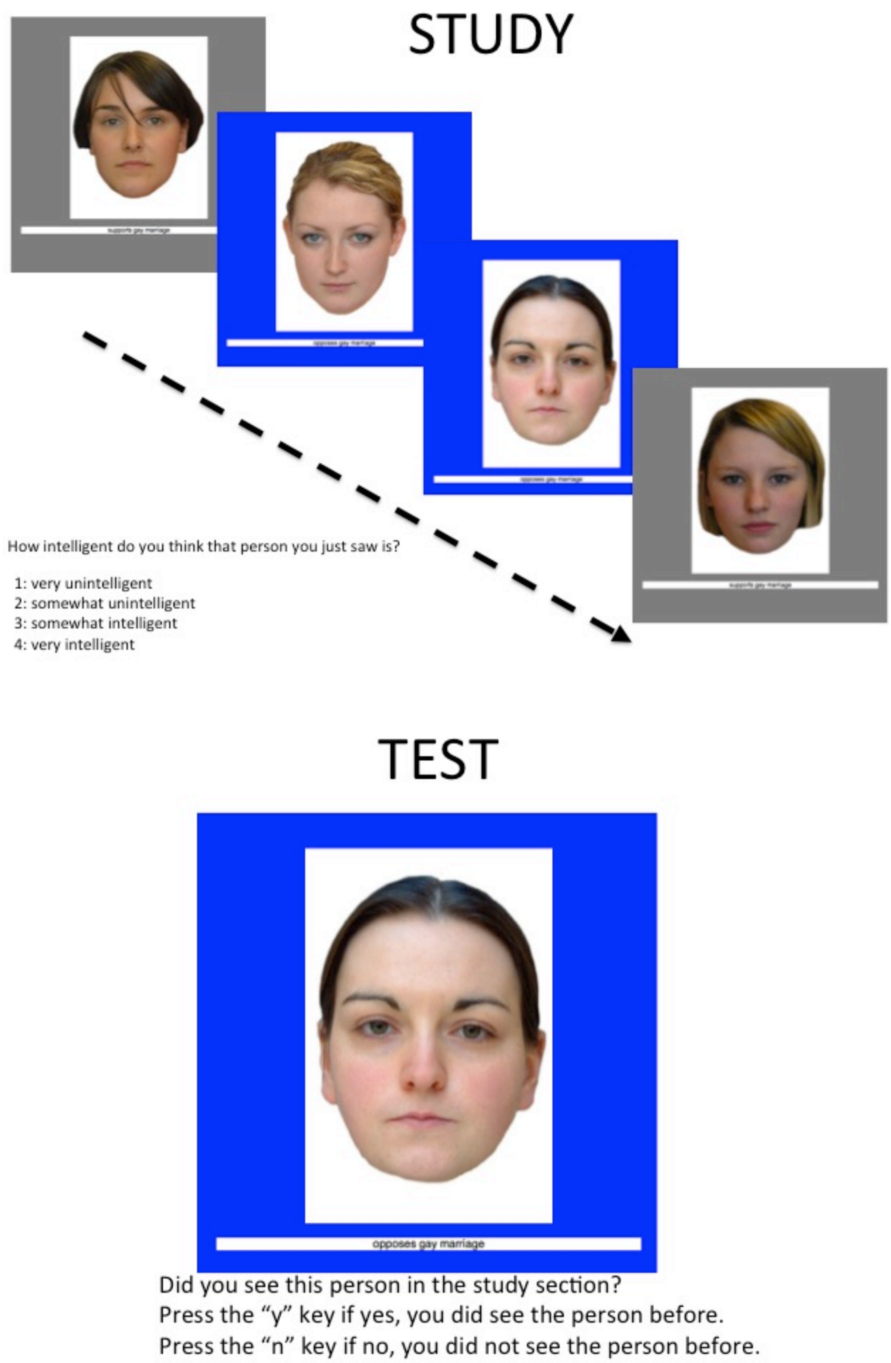


Figure 7. General Procedure for Experiments 2a-2g. The example stimuli are from Experiment 2b with the group factor of “supports” or “opposes” gay marriage.

intelligent, 4=very intelligent). They were also instructed that one color background (blue or grey) corresponded to one type of label that appeared with each face (e.g., “20 years old" and that the other color background corresponded to the other label (e.g., "30 years old").

The procedure for the study phase was as follows: stimuli were presented in random order with an ISI of 250 ms during which a fixation cross "+" was presented. Each face was presented onscreen for 3 s, then immediately following the face, participants were prompted to give an intelligence rating by pressing one of the keyboard keys 1-4.

After the study phase was complete, participants were given instructions for the recognition test. During this test phase, previously seen faces and previously unseen faces were presented one at a time and in a different random order for each participant. Previously seen faces were presented with their corresponding labels and background color. Half of the unseen faces were presented with the label "20 years old" and the rest with the label "30 years old". Each face remained onscreen until participants made a memory judgment response. Participants were to press the "y" button if they saw that person in the study phase or to press the "n" button if they did not see that person in the study phase.

RESULTS

The data for Experiments 2a and all ensuing experiments in this set are reported in Table 1. Paired-samples t-tests revealed no differences in d′ for memory performance or C for in-group (stimuli labeled as 20 years old) compared to out-group (stimuli labeled as 30 years old) faces, *p*s>.10.

Experiment 2b

Experiment 2b tested the effect of support for gay marriage on face recognition (supports vs opposes, again a novel factor that had not been previously tested).

METHOD

Participants. Thirty Caucasian female participants from AMT between ages 18-22 (mean age=20.6) received a payment of $2. None of the participants had participated in the two previous experiments.

Materials, Design, & Procedure. Experiment 2b was similar to Experiment 2a with one modification: the Group factor labels were "supports gay marriage" and "opposes gay marriage". In-group and out-group were analyzed based on whether or not participants reported support for gay marriage in a demographics survey.

RESULTS

Twenty-seven participants reported supporting gay marriage whereas three reported opposing it in the demographics survey. Therefore the in-group for the 27 participants supporting gay marriage was faces labeled with "supports gay marriage" and the in-group for the three participants opposing gay marriage was faces labeled with "opposes gay marriage". Paired-samples t-tests revealed no differences in d′ for memory performance or C for in-group compared to out-group faces, *p*s>.10.

Table 1. Mean d′ and C for Face Recognition as a Function of In-Group/Out-Group Factor for Online Experiments Using Amazon's Mechanical Turk (Experiments 2a-2c) and American University Participants (Experiments 2d-2g; in italics).

Experiment	In-Group		Out-Group		N
	d′	C	d′	C	
2a Psychological Age	1.10 (.52)	.06 (.28)	.97 (.46)	.09 (.39)	30
2b Gay Marriage	1.11 (.74)	-.56 (.37)	1.06 (.77)	-.53 (.39)	30
2c Occupation Status	1.23 (.55)	-.62 (.28)	1.18 (.53)	-.60 (.27)	30
2d Psychological Age	*1.70 (.71)*	*.16 (.30)*	*1.71 (.57)*	*.22 (.21)*	*10*
2e Gay Marriage	*1.33 (.55)*	*.10 (.39)*	*1.25 (.46)*	*.18 (.47)*	*19*
2f Physical Age	***1.63 (.53)*****	*-.07 (.27)*	***1.21 (.51)*****	*-.05 (.43)*	*30*
2g Gender	***1.56 (.58)****	*.15 (.47)*	***1.28 (.63)****	*.06 (.30)*	*30*

Note. **p*<.05, ***p*<.001. Standard deviations are in parentheses. For Occupation Status, High-Status is listed as In-Group and Low-Status is listed as Out-Group.

Experiment 2c

Experiment 2c tested the effect of occupation status (high vs low), a factor that had indeed led to an in-group bias previously (Ratcliff et al., 2011) but one that was not replicated in Experiment 1 of the present study.

METHOD

Participants. Thirty Caucasian female participants from AMT between ages 19-22 (mean age=20.8) received a payment of $2. None of the participants had participated in any of the previous experiments.

Materials, Design, & Procedure. Experiment 2c was similar to the previous experiments except that the Group factor of interest was occupation status. The same labels were used as in Experiment 1.

RESULTS

Paired-samples t-tests revealed no differences in d′ for memory performance or C for in-group (stimuli labeled with high-status occupations) compared to out-group (stimuli labeled with low-status occupations) faces, *p*s>.10.

Experiment 2d

Experiment 2d replicated 2a using American University participants and tested the effect of psychological age (20 years old vs 30 years old). The purpose was to see if somehow the participant population of AMT was not taking the task seriously as would a typical college participant population.

METHOD

Participants. Ten Caucasian female participants from American University between the ages 18-24 (mean age=20.8) received either an entry in a raffle to win a \$50 gift card or ½ a research credit for their Psychology courses. None of the participants had participated in any of the previous experiments. The experiment was otherwise identical to Experiment 2a.

RESULTS

Paired-samples t-tests revealed no differences in d′ for memory performance or C for in-group (stimuli labeled as 20 years old) compared to out-group (stimuli labeled as 30 years old) faces, *p*s>.10.

Experiment 2e

Experiment 2e replicated Experiment 2b using American University participants and tested the effect of support for gay marriage (supports vs opposes).

METHOD

Participants. Nineteen Caucasian female participants from American University between the ages 18-22 (mean age=19.7) received either an entry in a raffle to win a \$50 gift card or ½ a research credit for their Psychology courses. None of the participants had participated in any of the previous experiments. The experiment was otherwise identical to Experiment 2b in every aspect.

RESULTS

Paired-samples t-tests revealed no differences in d′ for memory performance or C for in-group compared to out-group faces, *p*s>.10.

Experiment 2f

In order to compare the effects of psychological/social factors to a more commonly tested physical factor, Experiment 2f tested physical age, where the in-group comprised faces of young adults and the out-group comprised faces of older adults.

METHOD

Participants. Thirty Caucasian female participants from American University between the ages 18-22 (mean age=19.5) received either an entry in a raffle to win a $50 gift card or ½ a research credit for their Psychology courses. None of the participants had participated in any of the previous experiments.

Materials, Design, & Procedure. Experiment 2f was similar to the previous experiments with the following modifications: 1) half of the stimuli were young females and the other half were older females (ages 64 and over) collected from FACES Database (Ebner, Riediger, & Lindenberger, 2010), and Ebner (2008), and 2) since the age grouping (own-age or older) of the stimuli was apparent, we did not use colored background or labels to indicate group membership.

RESULTS

Paired-samples t-test revealed a difference in d′ for memory performance, $t(29)=3.7$, $p<.001$, Cohen's d=.81, but no difference in C, $p>.10$. Participants were better at remembering own-age compared to other-age faces.

Experiment 2g

Experiment 2g tested gender, another commonly examined physical factor, to extend the comparison of psychological/social and physical factors. The stimuli were pictures of both young females and young males and, likewise, the participants were young females and young males.

METHOD

Participants. Thirty participants (18 female, 12 male) from American University between the ages 18-22 (mean age=19.8) received either an entry in a raffle to win a $50 gift card or ½ a research credit for their Psychology courses. None of the participants had participated in any of the previous experiments.

Materials, Design, & Procedure. Experiment 2g was similar to experiment 2f with the following modification: half of the stimuli were young females and the other half were young males collected from the Glasgow Unfamiliar Face Database (Burton et al., 2010).

RESULTS

Paired-samples t-test revealed a difference in d′ for memory performance, $t(29)=2.5$, $p<.05$, Cohen's d=.46, but no difference in C, $p>.10$. Participants were better at remembering own-gender compared to other-gender faces.

EXPERIMENTS 2a-2g CONCLUSIONS

We did not find any evidence for an in-group/out-group recognition bias for any of the psychological/social factors. This was inconsistent with the results of Ratcliff et al. (2011) who reported better memory performance for high-status compared to low-status occupations across three separate experiments. We only observed an in-group bias when a physical factor was the basis of the distinction between in-group vs. out-group. We observed an own-age bias in Experiment 2f and an own-gender bias in Experiment 2g. Thus, the psychological/social group factors do not appear to be as robust as physical group factors and may require special circumstances in order to emerge. With these findings in mind, in the next experiment we replicated Experiment 1 but used a physical factor that influenced the expected recognition bias.

CHAPTER 4

EXPERIMENT 3

Experiment 3 was similar to Experiment 1 with the following modifications. First, a physical factor of face age (own-young vs other-older) was manipulated such that the levels of eye region fan (low and high) were crossed with the levels of Group (in-group and out-group). Again, CIM predicts an effect of the factor of Group such that in-group faces should be better remembered than out-group faces but does not make a prediction about how featural fan would affect performance. Horry and Wright (2008) and Horry et al. (2010) found that context (background pictures) was more likely to be remembered with in-group than out-group faces, suggesting that a fan effect might emerge for own-age but not other-age faces.

METHOD

Participants. Sixty right-handed female participants from American University between ages 18-22 (mean age=19.3) received 2 extra credits for their Psychology courses or payment of $20. None of the participants had participated in any of the previous experiments.

Materials, Design, & Procedures. The materials and design were similar to Experiment 1 but with the following modifications (see Figures 8-10 for examples of the new stimuli and experimental design). The psychological/social factor of occupations was eliminated and replaced with the physical factor of age. Half of the stimuli consisted of the young female faces used in Experiment 1, and half consisted of older female faces taken from Ebner (2008) as well as Flickr.com, which were then photomanipulated according to the same specifications in Experiment 1. EEG/ERP (electroencephalogram/event-related potential) data were collected as part of a larger study, but only the behavioral results are reported here.

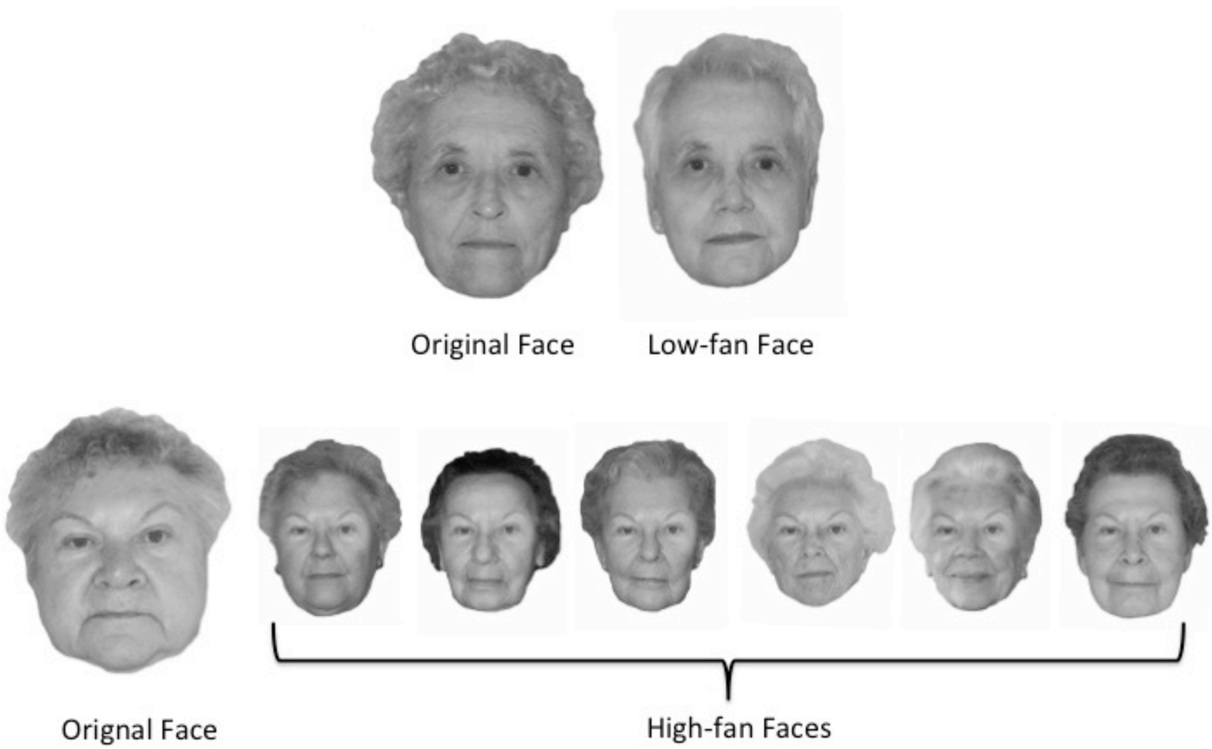


Figure 8. Examples of Low- and High-fan Older Adult Faces.

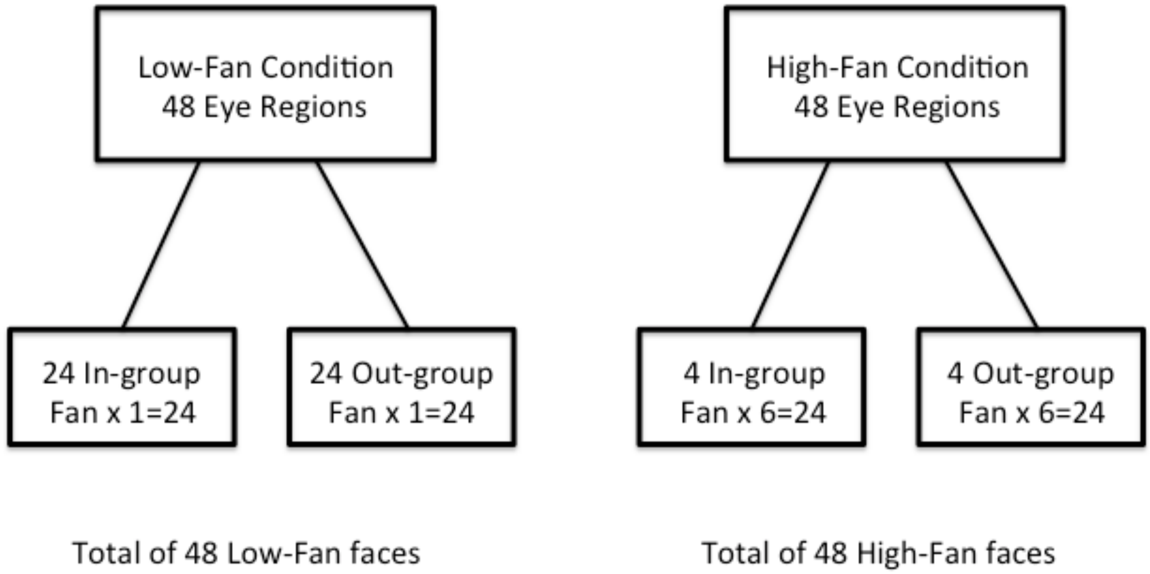


Figure 9. Design of the Study Phase for Experiment 3.

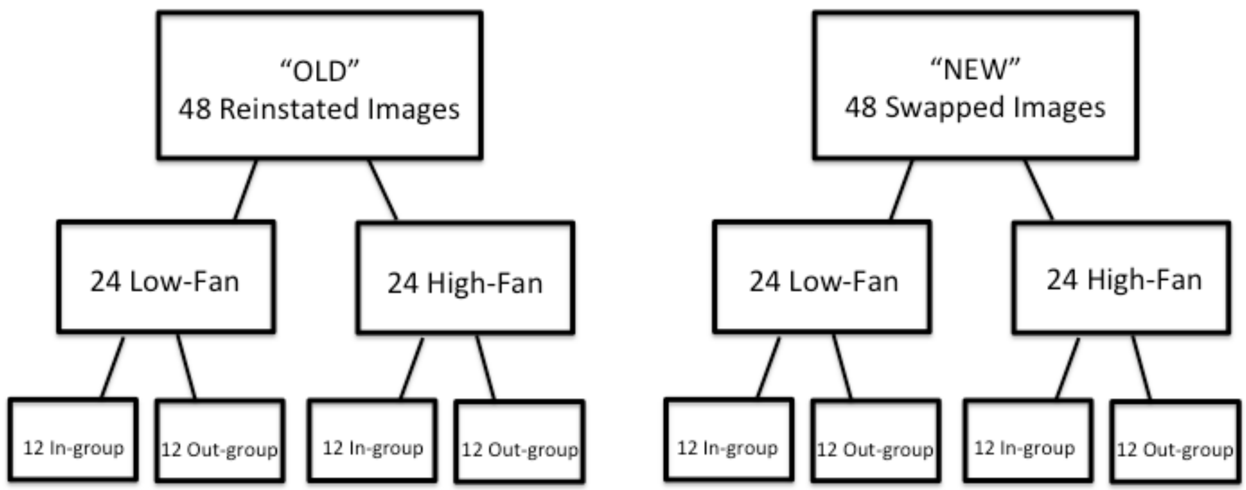


Figure 10. Design of the Test Phase for Experiment 3.

In order to ensure that there would be more low-fan observations for the EEG/ERP data, the design was modified such that there were an equal number of low-fan and high-fan face composites. Forty-eight faces were assigned to the low-fan condition and 48 were assigned to the high-fan condition. In each fan type, half were young faces and the other half were older faces. Thus, there were 24 unique face composites in the low-fan condition for each of the two face age groups. For the high-fan condition, for each of the two face age groups, four eye regions were randomly chosen for each participant and displayed with six faces for a total of 24 high-fan face composites. Overall, there were 96 stimuli in the study phase. At test, half of the faces were reinstated and half of the faces were swapped, just as in Experiment 1.

RESULTS

D′ results. The data are presented in Figure 11. A fan type (low vs high) x age (young vs older) x race (same vs other) repeated measures ANOVA revealed no effect of race, and no interactions of race with fan type or age, all Fs<1. Thus, we removed race from the model, and a fan type (low vs high) x age (young vs older) repeated measures ANOVA revealed a main effect of fan, $F(1, 59)=26.9$, $p<.001$, partial $\eta^2=.31$, and a main effect of age, ($F(1, 59)=11.1$, $p<.01$, partial $\eta^2=.16$. The interaction term between fan type and age was not significant, $F(1, 59)=1.6$, $p>.10$.

C results. Again similar results were obtained with the omnibus ANOVA and we removed race from the model to conduct a fan type (low vs high) x age (young vs older) repeated measures ANOVA. There were main effects of fan, $F(1, 59)=137.4$, $p<.001$, partial $\eta^2=.70$ and age, $F(1, 59)=76.7$, $p<.001$, partial $\eta^2=.57$. The interaction term between fan type and age was not significant, $F<1$.

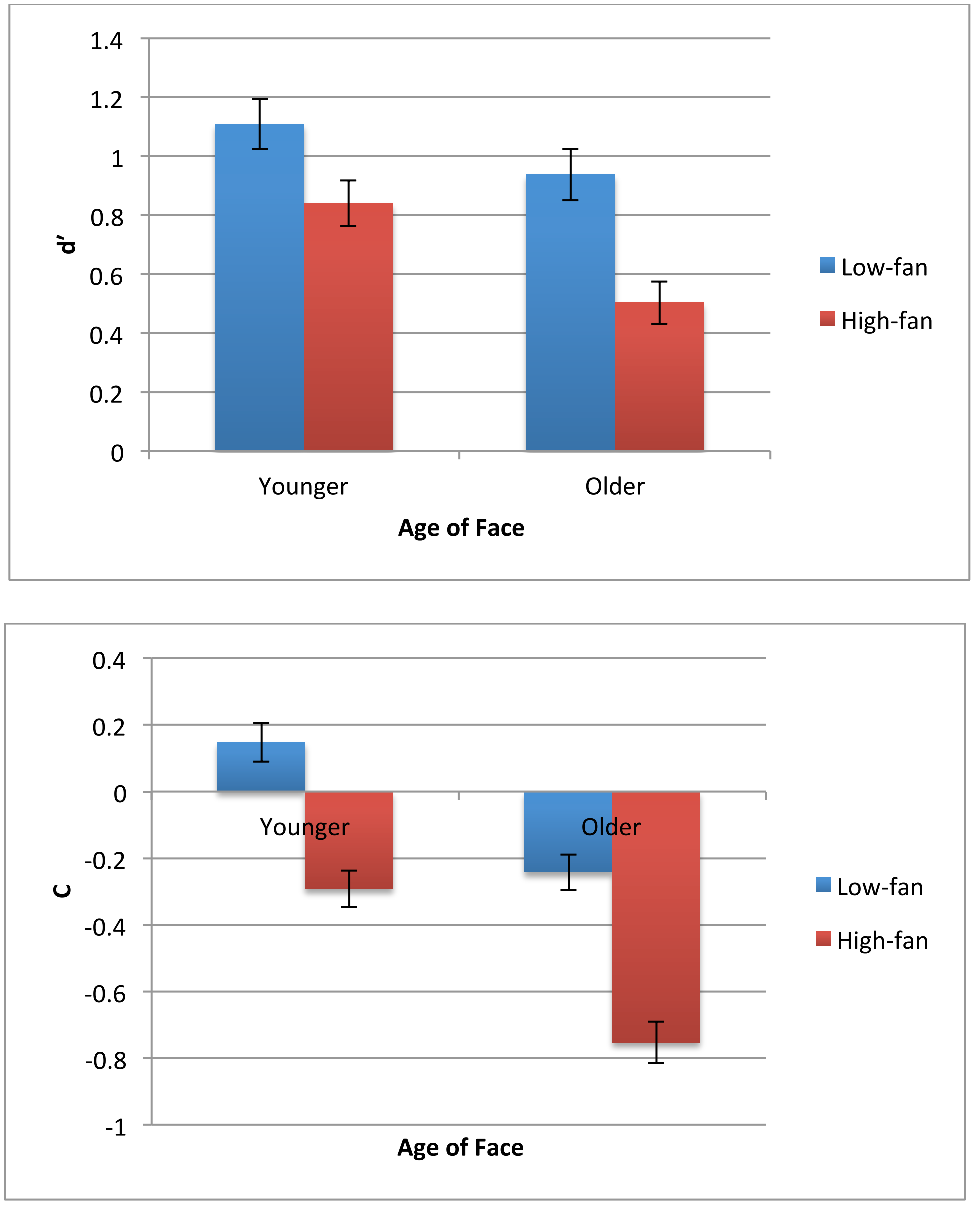


Figure 11. Experiment 3, Mean d′ and C as a Function of Age of Face and Fan Type. Error bars represent the standard error of the means.

For d′, there was an effect of fan such that low-fan faces (faces whose eyes were unique to one face) were better remembered than high-fan faces (faces that had the same eyes across six faces). There was also an effect of age such that young (own-age) faces were better remembered than older (other-age). There was however, no interaction between fan type and age indicating that relative distinctiveness did not interact with the own-age bias. The effect of fan type on own-age faces and other-age faces was equivalent. For C, there was an effect of fan such that participants were more likely to respond “Yes” to high-fan faces than low-fan faces. This high-fan liberal responding was also observed in Experiment 1, suggesting that at least for faces, high-fan of the eye region may induce liberal responding. There was an effect of age such that participants were more likely to respond “Yes” to older faces than young faces. Similar to d′, there was no interaction between fan type and age.

CHAPTER 5

GENERAL DISCUSSION

The aim of this study was to test face recognition performance predictions based on group membership (in-group/out-group) as well as to examine the role that relative distinctiveness (represented by a Featural Fan Effect) might play in face recognition in general and in face recognition biases in particular. In Experiment 1, we tested the Featural Fan Effect of the eye region and the psychological/social group factor of occupation status (high vs low status). We found an effect of fan such that low-fan faces were better remembered than high-fan faces, thus extending the work of Anderson and Paulson (1978) to photographs of real faces. A similar effect of fan emerged in Experiment 3 as well. However, there was no effect of occupation status as evidenced by our finding that memory performance was not better for high-status occupations compared to low-status occupations. That finding was inconsistent with both the predictions of the CIM model and the supporting results reported by Ratcliff et al. (2011). Because other research supporting CIM has also demonstrated group effects for psychological/social factors (Hugenberg & Corneille, 2009; Hugenberg et al., 2010; Hugenberg, Wilson, See, & Young, 2013), in Experiments 2a-2g we tested four psychological/social factors and compared them to two physical factors—age (own-age bias) and gender (own-gender effect) to explore the robustness of psychological/social factors in general. We found no effects for the psychological/social factors, although healthy effects emerged for physical factors. At this point, even though more controlled studies are needed to compare these two types of factors, these results suggest that perhaps some special conditions are required in order for psychological/social factors to elicit differential memory performance as a function of group membership.

In addition, in both Experiments 1 and 3, there emerged a difference in C such that participants were more likely to respond “Yes” to high-fan faces. Evidence from word font-fan literature also indicates that a liberal criterion is associated with high-fan fonts (Nyhus & Curran, 2009). There was also a similar tendency to respond more liberally to the out-group (older) faces in Experiment 3. High-fan faces share features and therefore look similar; correspondingly, physical out-group faces are not as frequently encountered as in-group faces and therefore are not a part of the face norm/template. It stands to reason that participants might be more likely to respond “Yes” to both because all of “those” faces looked the same. Other studies reporting within-experiment criterion shifts similar to ours occur when stimuli are more difficult to encode (Hockley & Caron, 2007; Hockley & Niewiadomski, 2007). Hockley and Caron reported a liberal responding to stimuli that were weakly encoded (given less study time) compared to more conservative responding to stimuli that were strongly encoded (given more study time). Hockley and Niewiadomski also reported a criterion shift according to the type of stimulus. Participants used a more liberal criterion to respond to word pairs compared to single words. In both studies, a liberal response criterion was associated with worse memory performance.

In general, in regards to the major face recognition bias models, the present results do not appear to support CIM, the EBH Model, or Sporer’s In-Group/Out-Group Model. Perhaps the current parameters of CIM might allow for a criterion shift in responding to out-group faces, such that a failure to individuate faces might lead to liberal responding. However, the lack of an in-group effect with psychological/social factors, the emergence of a Featural Fan Effect underscoring the role of distinctiveness, and a criterion shift as a function of the fan factor remove CIM from the class of viable models that can explain the current results.

The EBH model can account for our findings that physical factors (i.e. gender and age) affect recognition biases. However it is not clear how it could account for distinctiveness not based on atypicality (how unusual a face appears) a factor that was controlled for in our manipulations. EBH also does not appear to be able to account for a criterion shift. Likewise, Sporer's model also can account for biases based on physical factors, and, moreover, it has a provision for a criterion shift—more liberal responding to out-group faces which Sporer grounds in real-world application that a person may be more likely to identify an other-race person as the perpetrator of a crime. However, it cannot account for distinctiveness as specified by the Featural Fan Effect. As with the EBH model, only atypical faces would be more distinctive and therefore more memorable than typical faces. Thus, neither model can offer an explanation for why low-fan faces are remembered better than high-fan faces, since fan is an orthogonal manipulation on all typical faces.

Among the other models discussed, it appears that Valentine's Face Space Model (Valentine, 1991) would fit our data best. It can explain not only the recognition biases based on physical factors, namely the own-age effect in Experiments 2f and 3, and the own-gender effect in Experiment 2g (and be consistent with the lack of similar biases with psychological/social factors), but relative distinctiveness as manifested in the Featural Fan Effect as well. As described earlier, the model consists of an n-dimensional graphical depiction of “face space” where the most typical faces are spread around the origin, and less typical faces are spread out further away from the origin. Each face is graphed as a point amongst a space of n-dimensions, which, although unspecified, could theoretically represent age, gender, race, as well as types of features (e.g., shape of eyes). Given this, our same-age and same-gender faces would be spread around the origin whereas the other-age and other-gender faces would be clustered at a distance away from the origin. The “other” faces would be located within a restricted range and because expertise with those “other” non-exemplar faces does not allow for a wide range of discrimination based on the feature dimensions, the faces would tend to be grouped closely together. According to the model, faces spread around the origin would be easier to recognize than faces clustered together at a distance from the origin, and an own-age and own-gender bias would emerge.

Further, according to the Face Space Model, distinctiveness is operationalized as the density of face space that any given face occupies. If a given face is surrounded by many neighbors, or many faces that look similar based on one or more of the n-dimensions (e.g., race, age, shape of eyes, etc.), then it is said to be relatively typical. If a given face is not surrounded by many neighbors, then it is said to be relatively distinctive. Previously, Valentine and colleagues have examined distinctiveness in terms of how atypical a face appears to participants (e.g., is rated as atypical because of a prominent nose; Valentine, 1991, Valentine & Endo, 1992). Based on the n-dimension (e.g., nose shape) a distinctive face would be farther away from the origin than the typical faces. It would not be surrounded by many other faces on that particular n-dimension but would still be a part of the in-group (e.g., same- age, gender, race, etc). Instead of using typical and atypical faces, we created relative distinctiveness by manipulating the fan of the eye region. The model can still account for the better performance for the more distinctive faces that were rendered by being presented in the low-fan condition (see Figure 12).

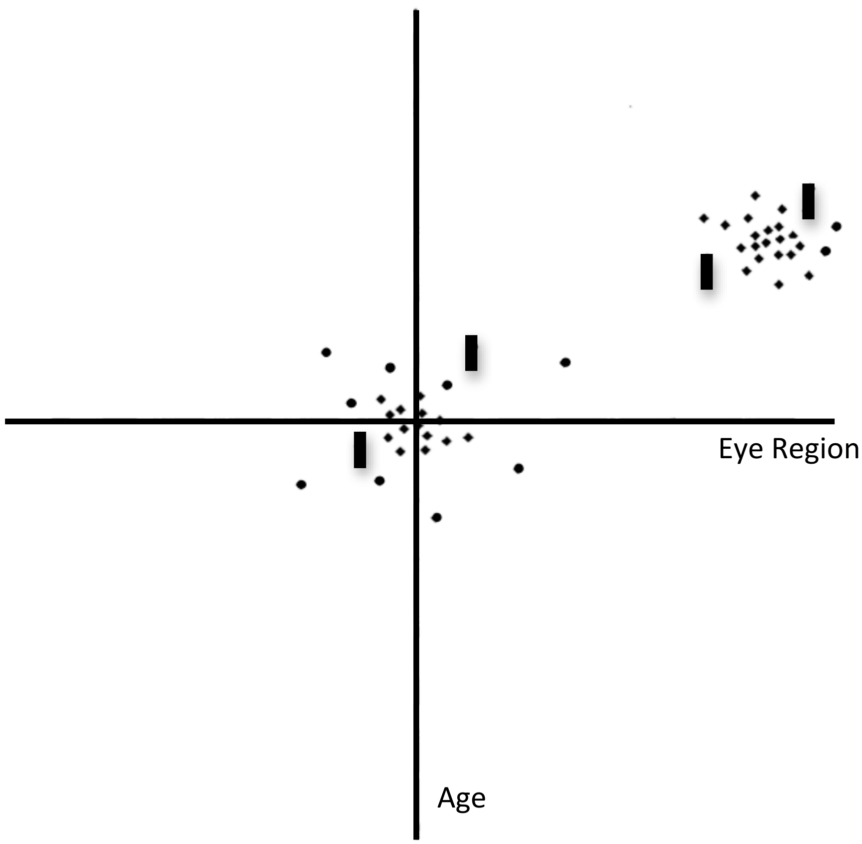


Figure 12. An Example of How the Face Space Model Might Account for Experiment 3. Faces are graphed on 2-dimensions depicting age and eye region. Single data points represent low-fan faces whereas clusters represent high-fan faces. Figure adapted from Towards an Exemplar Model of Face Processing: The Effects of Race and Distinctiveness, Valentine and Endo (1992), The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, Copyright © The Experimental Psychology Society, reprinted by permission of (Taylor & Francis Ltd, www.tandfonline.com on behalf of The Experimental Psychology Society.

That is, given that we manipulated the fan of the eye region and not how unusual a face appeared, all of our face stimuli would be classified as “typical” and should be equally spread around the origin. Then, the faces that shared eye regions (high-fan) would be tightly packed sub-clusters, rendering those high-fan faces more difficult to discriminate, since all share the same looking eye-region. Low fan-faces, which are similarly “typical”, would be spread around the origin but not tightly packed into sub-clusters like the high-fan faces. Therefore, the low-fan faces, as represented in the Face Space, should be easier to discriminate in a recognition test and our data support this. In summary, although other face recognition bias models can also account for relative distinctiveness as a function of atypicality, the Face Space can include differences in distinctiveness due to other means, such as the Featural Fan Effect.

An interesting finding in Experiment 3 was that there was no interaction between the Featural Fan Effect and the recognition bias created via the age manipulation. That is, making the in-group/out-group faces more or less relatively distinctive had no impact on the recognition bias itself. This finding is similar to that reported by Valentine and Endo (1992) who also did not find an interaction between the own-race bias and distinctiveness (although in that case distinctiveness was operationalized as unusual-looking faces). A possible explanation for lack of an interaction is that the mechanisms behind face recognition biases (at least with physical factors) and those that give rise to the Featural Fan Effect may be different and that the two tasks may be invoking the use of different types of information in the stimuli. For instance, it could be the case that, whereas same- and other- group faces differ by whether or not they are processed holistically or piecemeal, respectively, sharing (or not sharing) eye regions may have no impact on those types of face processing. Indeed, why should in-group faces that look similar or dissimilar be processed any more or less as Gestalts, or rather processed anymore more or less holistically? The same should be true of out-group faces—why should the Featural Fan Effect make them be processed any more or less like a set of features? The curiosity is that the fan manipulation worked in the same manner (at least behaviorally) on two supposedly different modes of face processing, holistic and piecemeal. Even though we found an own-age bias, the difference in memory performance for own-age vs other-age faces cannot rule out the possibility that all the faces, regardless of group membership, were processed in the same manner (e.g., all holistically or all piecemeal). For example, the own-age bias could have been driven by greater expertise with the eye region of one's own age group. Examination of how faces manipulated by Group and fan type are actually processed is something for future consideration.

Another consideration in regards to the Featural Fan Effect is how processing for faces might differ from that for words and fonts, which traditionally has been the domain in which the Featural Fan Effect has been studied. Although the behavioral patterns for word font fan and eye region fan might prove to be equivalent (low-fan items remembered better than high-fan items), the type of relative distinctiveness may be different. Specifically, changing the font of a word does not typically change the concept; however, changing the feature of a face changes the identity of the face. Indeed, there is evidence to indicate that when internal (eyes, nose, and mouth) or external (hair, chin, face outline) features are changed, participants respond neurologically as if the face were a new person (Andrews, Davies-Thompson, Kingstone, & Young, 2010). Therefore, words that share fonts would retain their concepts (the meaning of the words), but faces sharing features might not retain their unique identities.

A study by Wilford and Wells (2010) lends support to the notion that the Featural Fan Effect might work differently for faces and for other types of items. They reported that participants were faster to identify that a change had occurred in faces than houses, but were faster to localize that change in houses than faces. In other words, the assumption is that a face Gestalt prevents piecemeal processing and hurts identification of a featural change. This is certainly not the case with words because word processing does not hinder perception of changes in font, or rather, it is not difficult to localize changes in appearance of a word to the type of font.

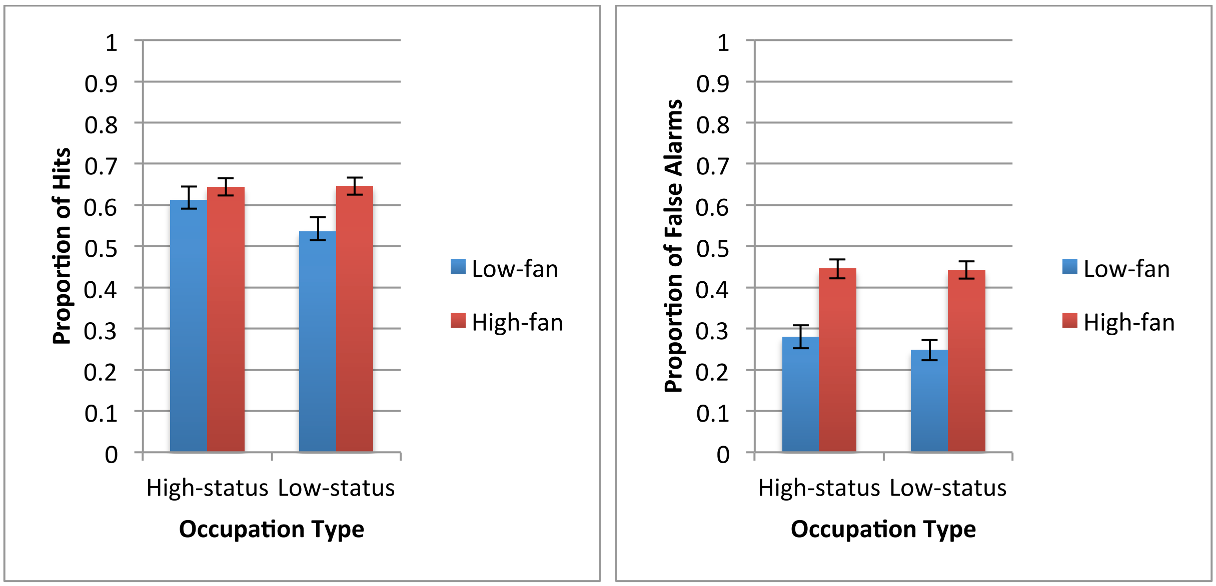
Finally, one secondary finding in Experiments 1 and 3 was that an own-race bias did not emerge. There are several reasons why an own-race bias would have been relatively weak in our experiments. Other-race participants were not of a single “other-race” and comprised several different races and ethnicities (e.g., African, Asian, Hispanic etc.) and some participants checked “other” because they were of more than one race (e.g., half Caucasian and half Asian). Therefore, “other-race” is really a melting pot of several races (and combinations of races) and not a homogeneous group as, for example, a sample of Asians from Korea (Michel et al., 2006). Further, although the own-race bias has been demonstrated with non-Caucasian races, Caucasians typically show a greater own-race bias than other races (Bruce & Young, 2012). Because we did not use non-Caucasian stimuli, we would have only elicited the weaker own-race bias of non-Caucasians.

In conclusion, the current work offers support for face recognition biases based on Group membership for physical factors (e.g., age) but not psychological/social factors (e.g., occupation status). In addition, the findings extend the Featural Fan Effect found with IDENTI-KIT faces (Anderson & Paulson, 1978) to real photographs. Perhaps more importantly, however, such relative distinctiveness created by manipulating the fan of the features does not appear to affect the own-age bias. The relative distinctiveness for faces created by the Featural Fan Effect does not rely on atypicality, and because of this, only one of the major face recognition biases models can account for the present results. Valentine's Face Space Model (1991) seems to be the best fit to not only explain the effects on face recognition of group membership but also relative distinctiveness in the guise of the Featural Fan Effect.

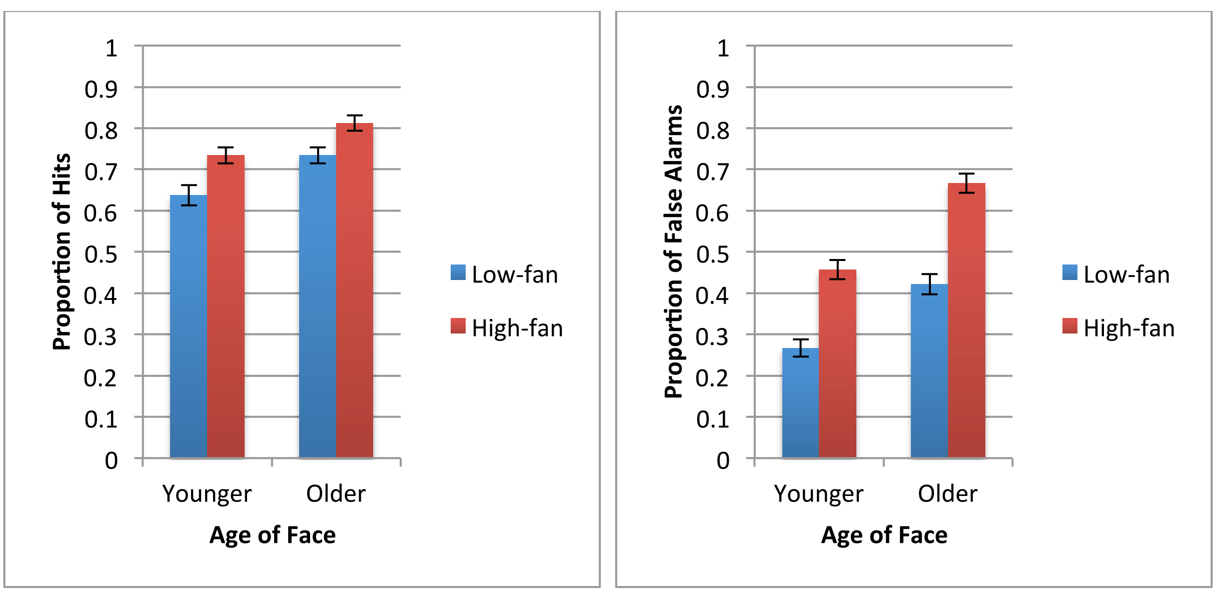
APPENDIX A

MEAN HITS, FALSE ALARMS (FAS) FOR FACE RECOGNITION AS A

FUNCTION OF IN-GROUP/OUT-GROUP AND FAN TYPE



Experiment 1



Experiment 3

APPENDIX B

MEAN HITS, FALSE ALARMS (FAS) FOR FACE RECOGNITION

AS A FUNCTION OF IN-GROUP/OUT-GROUP FACTOR

Experiment	In-Group		Out-Group	
	Hits	FAs	Hits	FAs
2a Psychological Age	.68 (.12)	.29 (.12)	.64 (.14)	.30 (.16)
2b Gay Marriage	.69 (.17)	.32 (.21)	.67 (.15)	.31 (.21)
2c Occupation Status	.67 (.15)	.25 (.16)	.66 (.16)	.27 (.17)
2d Psychological Age	*.73 (.16)*	*.17 (.09)*	*.72 (.12)*	*.15 (.08)*
2e Gay Marriage	***.71 (.12)****	*.25 (.15)*	***.66 (.15)****	*.24 (.17)*
2f Physical Age	***.79 (.11)****	*.25 (.10)*	***.72 (.15)****	*.30 (.15)*
2g Gender	*.72 (.16)*	***.21 (.16)****	*.70 (.15)*	***.27 (.13)****

Note. *p<.05. The data are from Experiments 2a-2g using Amazon's Mechanical Turk (Experiments 2a-2c) and American University Participants (Experiments 2d-2g; in italics). Standard deviations are in parentheses. For Occupation Status, High-Status is listed as In-Group and Low-Status is listed as Out-Group.

REFERENCES

Anderson, J. R., & Paulson, R. (1978). Interference in memory for pictorial information. *Cognitive Psychology*, *10*(2), 178-202.

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *The Journal of Neuroscience*, *30*(9), 3544-3552.

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition*, *10*(5), 527-536.

Bruce, V., & Young, A. (2012). *Face Perception*. Psychology Press: New York, NY.

Burton, A.M., White, D., and McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods*, *42*(1), 286-291.

Ebner, N. C. (2008). Age of face matters: Age-group differences in ratings of young and old faces. *Behavior Research Methods*, *40*(1), 130-136.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*(1), 351-362.

Caldara, R., Rossion, B., Bovet, P., & Hauert, C. A. (2004). Event-related potentials and time course of the 'other-race' face classification advantage. *Neuroreport*, *15*(5), 905-910.

Cross, J. F., Cross, J., & Daly, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. *Perception & Psychophysics*, *10*(6), 393-396.

Eimer, M. (1998). Does the face-specific N170 component reflect the activity of a specialized eye processor?. *Neuroreport*, *9*(13), 2945-2948.

Diana, R. A., Peterson, M. J., & Reder, L. M. (2004a). The role of spurious feature familiarity in recognition memory. *Psychonomic bulletin & review*, *11*(1), 150-156.

Diana, R. A., & Reder, L. M. (2004b). The other race effect: How do processing time and distracting contexts influence recognition of faces? Poster presentation at *Psychonomics*, Minneapolis.

Doi, H., Amamoto, T., Okishige, Y., Kato, M., & Shinohara, K. (2010). The own-sex effect in facial expression recognition. *NeuroReport*, *21*(8), 564-568.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, *13(*1), 8-20.

Going, M., & Read, J. D. (1974). Effects of uniqueness, sex of subject, and sex of photograph on facial recognition. *Perceptual and Motor Skills*, *39*(1), 109-110.

Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The perception of a face is no more than the sum of its parts. *Psychological Science*, *23*(4), 427-434.

Graf, P., & Ryan, L. (1990). Transfer-appropriate processing for implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 978-992.

Harrison, V., & Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic Bulletin & Review*, *16*(2), 264-269.

He, Y., Ebner, N. C., & Johnson, M. K. (2011). What predicts the own-age bias in face recognition memory?. *Social Cognition*, *29*(1), 97-109.

Hills, P. J. (2012). A developmental study of the own-age face recognition bias in children. *Developmental Psychology*, *48*(2), 499-508.

Hockley, W. E., & Caron, A. M. (2007, November). *Opposing strength- based mirror effects for words versus pictures: Evidence for within-list criterion changes*. Paper presented at the 48th Annual Meeting of the Psychonomic Society, Long Beach, CA.

Hockley, W. E., & Niewiadomski, M. W. (2007). Strength-based mirror effects in item and associative recognition: Evidence for within-list criterion changes. *Memory & Cognition*, *35*(4), 679-688.

Horry, R., & Wright, D. B. (2008). I know your face but not where I saw you: Context memory is impaired for other-race faces. *Psychonomic Bulletin & Review*, *15*(3), 610-614.

Horry, R., Wright, D. B., & Tredoux, C. G. (2010). Recognition and context memory for faces from own and other ethnic groups: A remember-know investigation. *Memory & Cognition*, *38*(2), 134-141.

Huang, I. N., & Wille, C. (1979). The von Restorff isolation effect in free recall. *The Journal of General Psychology*, *101*(1), 27-34.

Hugenberg, K., & Corneille, O. (2009). Holistic Processing Is Tuned for In-Group Faces. *Cognitive Science*, *33*(6), 1173-1181.

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.

Hugenberg, K., Wilson, J. P., See, P. E., & Young, S. G. (2013). Towards a synthetic model of own group biases in face memory. *Visual Cognition*, *21*(9-10), 1392-1417.

Hunt, R. R. (2013). Precision in memory through distinctive processing. *Current Directions in Psychological Science*, *22*(1), 10-15.

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, *6*(10), 766-774.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy evidence of perceptual narrowing. *Psychological Science*, *18*(12), 1084-1089.

Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(4), 811.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *The Quarterly Journal of Experimental Psychology: Section A*, *51*(3), 449-473.

Letourneau, S. M., & Mitchell, T. V. (2008). Behavioral and ERP measures of holistic face processing in a composite task. *Brain and Cognition*, *67*(2), 234-245.

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 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*(4), 559-574.

Lovén, J., Herlitz, A., & Rehnman, J. (2011). Women's own-gender bias in face recognition memory: The role of attention at encoding. *Experimental Psychology*, *58*(4), 333-340.

Macchi Cassia, V., Picozzi, M., Kuefner, D., & Casati, M. (2009). Why mix-ups don't happen in the nursery: Evidence for an experience-based interpretation of the other-age effect. *The Quarterly Journal of Experimental Psychology*, *62*(6), 1099-1107.

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law*, *7*(1), 98-118.

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.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual review of psychology*, *32*(1), 89-115.

Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, *17*(7), 608-615.

Neath, I., & Brown, G. D. (2006). SIMPLE: Further applications of a local distinctiveness model of memory. *Psychology of Learning and Motivation*, *46*, 201-243.

Nyhus, E., & Curran, T. (2009). Semantic and perceptual effects on recognition memory: Evidence from ERP. *Brain Research*, *1283*, 102-114.

Park, H., Arndt, J., & Reder, L. M. (2006). A contextual interference account of distinctiveness effects in recognition. *Memory & Cognition*, *34*(4), 743-751.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience*, *12*(5), 729-738.

Reder, L. M., Donavos, D. K., & Erickson, M. A. (2002). Perceptual match effects in direct tests of memory: The role of contextual fan. *Memory & Cognition*, *30*(2), 312-323.

Reder, L. M., Victoria, L. W., Manelis, A., Oates, J. M., Dutcher, J. M., Bates, J. T., ... & Gyulai, F. (2013). Why It's Easier to Remember Seeing a Face We Already Know Than One We Don't Preexisting Memory Representations Facilitate Memory Formation. *Psychological science*, *24*(3), 363-372.

Rhodes, G., Brake, S., Taylor, K., & Tan, S. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, *80*(3), 313-331.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review*, *13*(3), 499-505.

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.

Rossion, B., & Michel, C. (2011). An Experience-Based Holistic Account. *Oxford Handbook of Face Perception*, 215-243.

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

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46*(2), 225-245.

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*(1), B1-B9.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9*(4), 483-484.

Toseeb, U., Keeble, D. R., & Bryant, E. J. (2012). The significance of hair for face recognition. *PloS one*, *7*(3), e34144.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological review*, *80*(5), 352-373.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, *43*(2), 161-204.

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *The Quarterly Journal of Experimental Psychology*, *44*(4), 671-703.

Valentine, T., & Bruce, V. (1985). What's up? The Margaret Thatcher illusion revisited. *Perception*, *14*(4), 515-516.

Vizioli, L., Rousselet, G. A., & Caldara, R. (2010). Neural repetition suppression to identity is abolished by other-race faces. *Proceedings of the National Academy of Sciences*, *107*(46), 20081-20086.

Vokey, J. R., & Read, J. D. (1992). Familiarity, memorability, and the effect of typicality on the recognition of faces. *Memory & Cognition*, *20*(3), 291-302.

Von Restorff, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld. *Psychologische Forschung*, *18*(1), 299-342.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(4), 442-452.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., ... & Johnstone, T. (2004). Human amygdala responsivity to masked fearful eye whites. *Science*, *306*(5704), 2061-2061.

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.

Wilford, M. M., & Wells, G. L. (2010). Does facial processing prioritize change detection? Change blindness illustrates costs and benefits of holistic processing. *Psychological Science*, *21*(11), 1611-1615.

Wolff, N., Kemter, K., Schweinberger, S. R., & Wiese, H. (2014). What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias. *Social Cognitive and Affective Neuroscience*, *9*(5), 580-590.

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*, *114*(1), 101-114.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747-759.