Verifying other-race identity in forensic settings: Increasing tolerance for variability in appearance

by
Claire M. Matthews

A thesis submitted in partial fulfillment
of the requirements for the degree
Master of Arts

Department of Psychology
BROCK UNIVERSITY
St. Catharines, Ontario

June 2017

Abstract

When viewing unfamiliar faces, photos of the same person often are perceived as belonging to different people and photos of different people as belonging to the same person. Identity matching of unfamiliar faces is especially challenging when the photos are of a person whose ethnicity differs from that of the observer. In contrast, matching is trivial when viewing familiar faces, regardless of race. In a 1-in-30-lineup task in which participants are asked to find the image of a target from an array of 30 identities, viewing multiple images of an own-race target improves performance, reflecting rapid familiarization (Dowsett et al., 2016). Here, participants were asked to find an other-race target from an array of 30 images, and participants' performance on this task was observed as they were provided with additional images of each other-race target. I report rapid familiarization when the target was known to be present (Experiment 1) but not when both target-present and target-absent trials were included in the task (Experiment 2). Viewing multiple images of a target-absent identity provided no benefit in reducing false alarms. Although a possible route to familiarization with other-race faces, my findings suggest caution for the use of multiple images in applied face verification settings.

Acknowledgments

I wish to thank everyone who provided me with support throughout the completion of my Masters thesis. First and foremost, I would like to thank my supervisor, Dr. Cathy Mondloch. Thank you for your constant guidance, expertise and support throughout the completion of this project and for continuing to encourage me as I pursue my PhD. I would also like to thank my committee members Dr. Angela Evans and Dr. Karen Arnell for their valuable feedback throughout the completion of this project.

Thank you to all the members of the Face Perception Lab for all your help with this project and for the many laughs we shared along the way. Special thanks to Harmonie Chan for collecting the many images used as stimuli in this study and Gaby Salgado for all your help organizing and entering data. Thank you to my fellow graduate students for your support and for making this such an enjoyable process.

Finally, I would like to thank my friends and family for are their unconditional support throughout the last two years. Thank you Cameron, for always encouraging me to reach my goals and for your enthusiasm and interest in my research.

Table of Contents

Figures

Appendices

Verifying other-race identity in forensic settings: Increasing tolerance for variability in appearance

Matching faces to photographs is heavily relied upon in many applied settings. It is used as a method of identity verification when crossing international borders, obtaining official documents, and buying age-restricted goods. When matching identity, the perceiver is faced with two distinct challenges: discrimination and recognition. The perceiver needs to be able to discriminate between identities in order to detect instances where the photo ID does not match its bearer. Perceivers also need to be able to recognize an identity despite natural variability in appearance. This could involve detecting that the photo ID is valid even after several years of aging or changes in makeup and hairstyle. Despite technological advancements in facial recognition software, these systems are not yet entirely automatic and therefore, human observers are often still required to interpret the output and make the final matching decision in forensic face verification settings (White, Dunn, Schmid & Kemp, 2015). This reliance on human perceivers to verify identity by matching faces is problematic, as humans are highly error-prone at performing this task.

Checking photo identification may seem like an effective method of identity verification because matching faces to photographs seems like a trivial task. This is because we are experts at recognizing familiar faces (Bruce, 1982; Burton, Wilson, Cowan & Bruce, 1999). We can all easily tell our family members, friends or favourite celebrities apart from individuals who look similar to them, and we can also recognize these familiar people despite changes in their appearance. Familiar face recognition is robust and high accuracy persists despite distortions to the face and poor viewing conditions. Hole, George, Eaves, and Rasek (2002) compared participants' recognition accuracy of unaltered celebrity faces to distorted versions of these faces. Accuracy was not influenced by stretching the face vertically, and minimally influenced by stretching the face horizontally. Further, in a study by Burton et al. (1999), participants viewed video clips from closed-circuit television (CCTV) footage. The people captured on the CCTV footage were lecturers personally familiar to one group of participants and unfamiliar to the other. After viewing the videos, participants were shown high-quality images and asked to indicate whether each person had appeared in one of the videos. Participants unfamiliar with the people in the videos performed poorly on the recognition task. However, participants who were familiar with the people in the videos performed almost perfectly, despite the poor quality of the CCTV footage.

However, in applied settings, faces are typically unfamiliar and recognition of unfamiliar faces is highly error-prone (see Burton & Jenkins, 2011 for review). Error rates of up to 30% have consistently been reported in laboratory tasks that mimic real-world identity verification tasks. These error rates are found regardless of whether participants are asked to match an image to a nearly identical photo (Bruce et al., 1999; Bruce, Henderson, Newman & Burton, 2001; Burton, White, & McNeill, 2010; Megreya & Burton, 2006; 2008; Megreya, Sandford & Burton, 2013) or to a live person (Kemp, Towell, & Pike, 1997; Megreya & Burton, 2008). Kemp et al. (1997) conducted a study using real cashiers in a supermarket. At the time, the United Kingdom had just begun to introduce credit cards that included photographs of the bearer. Therefore, cashiers needed to be able to verify that the image on the credit card matched the person making a purchase in order to avoid credit card fraud. Undergraduate students were asked to act as real shoppers presenting their credit cards when making a purchase. For half of the transactions, the bearer matched the picture on their credit card, and for half they did not. Cashiers were asked to decide whether they would accept each credit card. Over 50% of fraudulent cards were falsely accepted and about 10% of valid cards were falsely rejected by cashiers, suggesting that even individuals highly motivated to detect fraud are poor at matching faces to photographs.

Surprisingly, even professionals trained to use photographs to verify identity on a daily basis are highly error-prone at matching unfamiliar faces. In a study by White, Kemp, Jenkins, Matheson, and Burton (2014) the matching ability of passport officers was compared to that of untrained undergraduate students. Both groups of participants viewed two images and were asked to complete the task faced by passport officers: to decide whether the two images were of the same person or of two different people. Participants also completed this task comparing a live person to an image. Across both tasks, high error rates were found, with passport officers demonstrating similar matching ability to that of untrained students. Further, passport officers' length of employment did not predict their face matching ability, suggesting that even highly experienced passport officers are highly error-prone. These results demonstrate the importance of investigating ways in which unfamiliar face matching can be improved in forensic face verification settings.

Recognition When Facial Appearance Varies Naturally

In the vast majority of studies examining unfamiliar face matching, the images used were tightly controlled (e.g., taken with the same camera, from the same distance, under identical lighting conditions and with neutral expressions). When unconstrained images that capture natural variability in appearance (e.g., changes in lighting, viewpoint, expression, makeup etc.) are used, error rates increase (Megreya et al., 2013; see Burton, 2013). Indeed, even modest variability—variability that results from taking two images under the same conditions but several months apart—increases error. Megreya and colleagues asked participants to find a target from a lineup of 10 images in which the target image and the correct image in the lineup were either captured on the same day or several (M= 17.2) months apart. Accuracy was significantly lower for images taken months apart than for images taken on the same day. Specifically, for images taken months apart, participants only correctly selected the target 58% of the time on target-present trials (i.e., hits) compared to 79% of the time for images taken on the same day. Further, misses (incorrectly indicating the target is missing from the lineup) and misidentifications (incorrectly selecting one of the distractors in the lineup) rose from 11.3% and 9.7% respectively for images taken on the same day to 27.7% and 13.7% respectively for images taken months apart. This significant increase in error rates was found despite the fact that these stimuli were still highly controlled (i.e., all captured from the same viewpoint with the same neutral expression).

Error rates increase further when using images that capture the type of natural variability in appearance that is encountered in the real word; the first examination of this was conducted by Jenkins, White, Van Montfort and Burton (2011). In this study, participants were asked to sort 40 ‘ambient images’ (images that incorporate unconstrained variability in appearance) into piles such that each pile included all of the photos of one identity. Participants were unaware that the 40 images belonged to only two people (20 images each). Participants familiar with the identities correctly made two piles; participants unfamiliar with the identities made an average of seven piles. This means that they perceived there to be seven different identities, instead of just two.

The impact of such errors is clearly evident in eyewitness identification, in which the suspect's appearance can vary significantly from the witness's memory. The Innocence Project is an organization formed to help exonerate individuals wrongfully convicted of crimes. Misidentification errors (i.e., identifying someone from the lineup who was not the perpetrator) contributed to the false incarceration of about 75% of over 300 people exonerated so far by The Innocence Project (Innocence Project, 2016). Therefore, it is imperative that we understand the process by which a face goes from being unfamiliar to familiar in order to inform protocols to facilitate accurate recognition of unfamiliar faces.

Difficulty in recognizing unfamiliar faces in ambient images results from unfamiliar face representations being based largely on lower-level image properties (reviewed in Hancock, Bruce & Burton, 2000; Young & Bruce, 2011). Even slight changes in image properties such as lighting conditions, viewpoint, facial expression and aesthetics make it difficult to recognize an unfamiliar face. This means that unfamiliar face representations are specific to the particular instance in which the image was captured or the conditions under which a new person was encountered and thus a change in image appearance is perceived as a change in identity. In contrast, familiar face representations are robust and unaffected by variation in appearance (see Burton, 2013). This suggests that building robust representations of newly encountered faces facilitates the process by which these faces become familiar.

Several studies suggest that although within-person variability in appearance makes unfamiliar face matching more challenging, it might also be the route by which a face becomes familiar (Andrews, Jenkins, Cursiter & Burton 2015; Bindemann & Sandford, 2011; Dowsett, Sandford & Burton, 2016; Menon, White & Kemp, 2015b; Ritchie & Burton, 2017). In Ritchie and Burton's (2017) study participants learned 10 images of 20 people; each set of images contained either high or low variability in appearance. In a subsequent recognition task, in which new images of these identities were presented, recognition accuracy was higher in the high-variability condition. In Dowsett and colleagues' (2016) study participants were asked to identify a target in a 30-image lineup. Participants were first provided with one image of the target. An additional photo was added on each trial until participants simultaneously viewed six images of the target identity. To investigate whether any learning was identity-specific, or transferred across identities, participants completed this task with three different target identities. Accuracy increased as images were added. Additionally, the observed learning was identity-specific; exposure to the ways in which one target identity varied did not improve recognition of subsequent identities. This is consistent with findings suggesting that within-person variability in appearance is idiosyncratic in nature (Young & Burton, in press). Collectively these studies show that abstract representations of an identity emerge as new exemplars of that identity are encountered (White, Burton et al., 2014).

That variability is a route to learning is consistent with developmental theory. In particular, variability seems to play a significant role in the perceptual development of face recognition, as well as spoken word recognition (i.e., in language development; Watson, Robbins & Best, 2014). Processing spoken language involves balancing the same two challenges associated with face perception: discrimination and recognition. For example, when we hear the word “bed” spoken, we need to be able to perceive that it is different from when the word “red” is spoken, despite the words sounding quite similar. We also need to be able to recognize the word “bed” across variability in the way that it is spoken (e.g., when spoken by a speaker with a different accent or tone of voice). Watson and colleagues (2014) suggest that a common principle of face and language recognition during development is the importance of exposure to variability during the process of perceptual narrowing. Perceptual narrowing occurs to tune the perceptual system to stimuli relevant to the environment (e.g., to faces of our own-race and speech sounds of our native language). Exposure to variability during this process tunes the system to better recognize relevant stimuli across natural variability in the future. For example, children are better able to learn new words when spoken by multiple different speakers (Watson et al., 2014). Therefore, exposure to variability may be crucial to optimal perceptual learning.

Variability has also been shown to provide optimal training of perceptual expertise with objects (Gonzalez & Madhavan, 2011; Hussain, Bennett, & Sekuler, 2012). In a study by Gonzalez & Madhavan (2011), participants were trained to detect potentially dangerous objects in one of two training conditions. Participants in the low variability group were trained on five dangerous items all from the same category (e.g., five knives), while participants in the high variability group were trained on five items from all different categories (e.g., a knife, a gun, etc.). Participants then completed a task meant to simulate luggage going through a security scanner in which they were asked to detect novel dangerous objects. Compared to participants in the low variability group, participants exposed to high variability training were significantly more accurate at identifying novel dangerous objects.

Recognition of Other-Race Faces When Facial Appearance Varies Naturally

Here, I investigated whether exposure to within-person variability in appearance might offer a partial solution to an especially poignant and all too common scenario: when errors are multiplied because the person making the identity decision is of a different ethnicity than the person in question. Across numerous paradigms other-race faces are remembered less well than own-race faces (Meissner & Brigham, 2001). This has been demonstrated in old/new recognition tasks in which participants study a series of faces and are later asked to recognize previously seen faces intermixed with novel faces (e.g., Golby, Gabrieli, Chiao, & Eberhardt, 2001; MacLin & Malpass, 2001; Wright, Boyd, & Tredoux, 2003), as well as in tasks in which participants are asked to identify a target in a lineup from memory (e.g., Evans, Marcon, & Meissner, 2009; Jackiw, Arbuthnott, Pfeifer, Marcon, & Meissner, 2008; Meissner, Tredoux, Parker, & MacLin, 2005). This phenomenon is also quite evident in the real world; in a disproportionate number of the cases in the Innocence Project, cross-race identification was a contributing factor to wrongful conviction (Innocence Project, 2016).

Although commonly posed as a problem of impaired memory for other-race faces, the other-race effect (ORE) is also evident when memory demands are limited. The ORE is evident in 1-to-1 matching tasks in which participants are presented with two images and asked to decide whether they belong to the same person or to two different people (Meissner, Susa & Ross, 2013; Mondloch et al., 2010). The ORE has also been demonstrated in lineup tasks in which the memory component is all but eliminated

(Megreya, White & Burton, 2011). For example, in a study by Megreya et al. (2011), UK and Egyptian participants were simultaneously presented with a target face and an array of 10 images and were asked to find that target in the lineup. This type of task minimizes memory demands, as participants can visually scan back and forth between the target image and the 10 images in the lineup. Both UK and Egyptian target identities were used, creating both own-and other race trials for all participants. The results demonstrated that even with low memory demands, the other-race effect persists; participants were more accurate at recognizing and correctly rejecting own-race targets compared to other-race targets. This suggests that the ORE may arise during the perceptual encoding of other-race faces, and therefore exposure to within-person variability at that stage may be beneficial.

The studies described above have characterized the ORE as a problem of discrimination. Using one or two tightly controlled images per identity precludes investigating an equally important challenge: recognizing the same person across changes in appearance. Recognizing identity in ambient images is especially challenging for other-race faces (Laurence, Zhou & Mondloch, 2016; also see Hayward, Favelle, Oxner, Chu & Lam, 2017). Laurence and colleagues (2016) used Jenkins et al's (2011) sorting task in which participants sort 40 ambient images into piles based on identity; they provided the first test with both own-and other-race faces. Participants made nearly twice as many piles when sorting other- compared to own-race faces. Interestingly, misidentification errors (i.e., failing to discriminate between identities) were rare. The errors most commonly made by participants were creating too many piles (i.e. failing to recognize the same person across natural variability in appearance), a challenge that has been largely ignored in past research using only one or two images to represent an identity. Although matching unfamiliar faces across variability in appearance is especially challenging when faces are unfamiliar, the ORE disappears when a face is highly familiar. Zhou & Mondloch (2016) asked Chinese participants to complete the sorting task with images of two Caucasian celebrities with whom they were quite familiar; participants sorted these faces perfectly into two piles. This raises questions about the types of training or experience that might improve recognition for other-race faces.

Emphasizing the ORE as a problem of discrimination has led to the development of training paradigms that are based on individuating tightly controlled faces that are highly similar to each other. Previous attempts at improving other-race face memory have produced only modest results (McGugin, Tanaka, Lebrecht, Tarr & Gauthier, 2011; Tanaka & Pierce, 2009). For example, in a study by Tanaka & Pierce (2009) participants were trained to individuate eight different unfamiliar other-race identities. After training participants showed some improvement on a recognition task. However, such training involved learning only one image per identity, failing to expose participants to the ways in which each other-race identity can vary. To my knowledge, no study to date has examined if exposure to within-person variability in appearance may be beneficial to improving other-race face recognition.

The Present Study

I provide the first examination of whether exposure to natural variability in appearance is a route to familiarization with other-race faces. In Experiment 1, I employed Dowsett and colleagues' (2016) method with other-race faces. Participants were gradually introduced to six different images of an other-race target identity and were asked find that identity from a 30-image array after each image was presented. Images were selected to capture natural variability in appearance. To investigate whether any learning was identity-specific or transferred across identities, each participant completed this task with three targets. To increase generalizability, the current study rotated through 30 different target identities, whereas Dowsett et al. (2016) used the same three target identities for all participants. I predicted that, as shown previously for own-race faces (Dowsett et al., 2016), there would be no transfer of learning across identities. Theoretically, this is consistent with evidence that within-person variability is idiosyncratic in nature (Young & Burton, in press). The ways in which Identity A vary are not the same as the ways in which Identity B vary, so learning the ways in which one identity varies might only be beneficial to improving recognition for that specific identity. Therefore, if improvement from exposure to within-person variability is identity specific, then I would expect accuracy after viewing only one image to be the same across the three identities, showing no transfer of learning.

In Experiment 2, I addressed a limitation common to both Dowsett et al.'s study and Experiment 1: including only target-present trials precludes false alarms. False alarms occur when participants incorrectly select a distractor from the lineup when the target is in fact absent. In Experiment 1, this type of mistake was not possible, as the participants were instructed to find a target that was definitely present in the lineup. In the context of eyewitness identification, the perpetrator is not always present in the lineup and wrongful convictions can occur if the true perpetrator is absent from the lineup but the witness selects an innocent suspect. Thus, in Experiment 2 I investigated whether within-person variability helps when target-absent trials are included. Specifically, I was interested in investigating whether exposure to multiple images enhances both recognition of the target under conditions of uncertainty on target-present trials and detection that the target is not in the lineup on target-absent trials. If exposure to within-person variability facilitates the formation of a robust representation that allows accurate recognition of the target when they are present in the lineup, and correct rejection of the distractors when the target is absent from the lineup, then I would expect accuracy to increase as additional images are presented to participants in both target-present and target-absent lineups. Alternatively, if exposure to multiple images is simply expanding tolerance for variability rather than forming a robust representation of the specific target, participants might not benefit from viewing multiple images when the target is absent.

Experiment 1: Method

Participants

Forty undergraduate students (33 female, M_{age}=19.35, Range=18-24) participated. This sample size was selected to be comparable to Dowsett et al. (2016) and allow counterbalancing of images. Participants completed a questionnaire assessing their face-to-face interaction with Black individuals (e.g., number of hours of contact per week, number of family members, friends, etc.). All participants reported minimal interaction with Black individuals; all reported having fewer than three Black friends, with 20 reporting no Black friends. An additional three participants were excluded: two because they expressed extensive experience with Black individuals and one because of experimenter error. All participants provided written informed consent and received research credit for their participation.

Stimuli

A total of 30 Black female identities were used as stimuli. These identities were a mixture of celebrities and non-celebrities obtained from the Internet via Google images. All identities were expected to be unfamiliar to participants (confirmed by the familiarity questionnaire). Seven images of each identity were collected in which the face exceeded 150 pixels in height, was free from occlusions and shown from a frontal view. The images were selected to capture natural variability in each identity's appearance (see Figure 1); images were not standardized or manipulated. The 30 identities were not specifically chosen to look alike. All images were cropped to 286 by 400 pixels at 72 pii and were printed in colour on 3.95 X 6 cm individual cards. One image of each identity was selected randomly to be placed in the 30-image array. The other six images of each identity were used as training images.

Procedure

Participants were first given 30 images laid out in a 5x6 array on the table in front of them. Participants were asked to find a target identity out of the lineup and were ensured that the target would in fact be present. The participants were then given the first image of the first identity, and were asked to make their selection. After recording their response, the chosen image was placed back into the 30-image deck, which was then shuffled and laid back out on the table. For the second trial, a second image of the first identity was given to the participant, with the first image remaining visible. Again,

Figure 1. A portrayal of the kind of variability encountered in the task. All six images depict the same person. Copyright precludes showing the images used in my study.

participants were asked to select the image they believed to be the target. This procedure continued until participants were able to view six different images of the first target identity, with the order in which images were presented randomized. The experimenter made it clear to the participants that each of the six images depicted the same person. Participants were told that they could change their response from trial to trial. No feedback was given at any time.

Each participant performed this task with three identities. Across participants, 30 different target identities were used; each identity served as a target for four different participants. The 30-image array did not change across participants or across trials. After completing the lineup tasks each participant completed a questionnaire assessing their familiarity with the identities used; all participants reported being unfamiliar with all faces.

Results

Accuracy

As shown in Figure 2, the percentage of participants who correctly identified the target from the lineup increased as more images of the target were presented, with no transfer across identities. Linear contrast analyses (as in Dowsett et al., 2016) confirmed that accuracy increased as more images of each target were added for the first two identities [Identity A, $F(1,39) = 5.17$, $p = .029$, $\eta_p^2 = .12$; Identity B, $F(1,39) = 11.092$, $p = .002$, $\eta_p^2 = .22$]. The results of the linear contrast for the last target identity approached significance, $p = .082$. Paired sample t-tests (1-tailed) comparing accuracy after viewing only one image to accuracy after viewing six different images confirmed that participants were more accurate at correctly choosing each target out of the lineup when they were

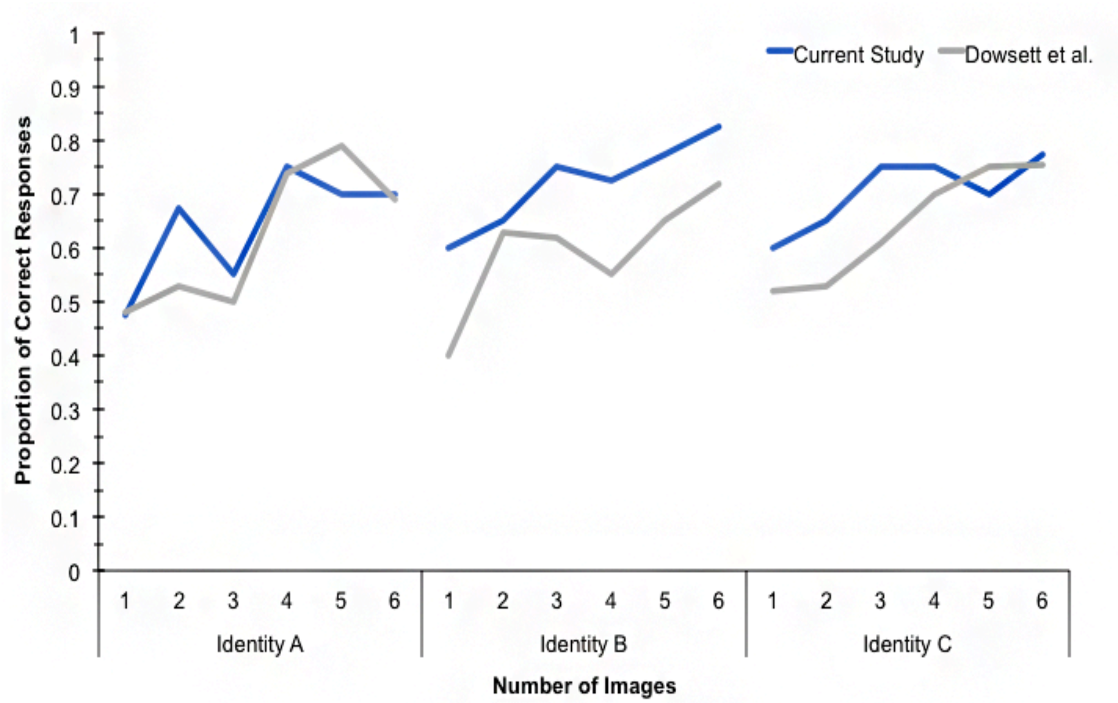

Figure 2. Proportion of participants who correctly selected the target. Grey lines depict own-race results by Dowsett et al., (2016).

able to view six images compared to just one [Identity A, 70.00% vs. 47.50%, $t(39) = 2.47$, $p = .009$, $d = 0.46$, Identity B, 82.50% vs. 60.00%, $t(39) = 2.97$, $p = .003$, $d = 0.51$, Identity C, 77.50% vs. 60.00%, $t(39) = 1.74$, $p = .045$, $d = 0.38$]. On average, participants selected about two different identities across the six trials for each target.

Transfer of Learning

Completing the task with multiple images of the first identity did not reliably improve performance for the next two identities. Paired sample t-tests (1-tailed) comparing performance on the first trial for the second and third identities to the first trial for the first identity revealed no transfer [Identity A Trial 1 (47.50%) vs. Identity B Trial 1 (60.00%), $p = .128$; Identity A Trial 1 vs. Identity C Trial 1 (60.00%), $p = .141$]. Paired sample t-tests comparing performance on the last trial for the second and third identities to the last trial for the first identity revealed improvement from the first (70.00%) to second (82.50%) identity, $t(39) = 1.71$, $p = .048$, $d = 0.29$, but not from the first to third (77.50%) identity, $p = .162$.

In summary, five of six measures indicated improved recognition as images were added but only one of four measures indicated improved recognition across identities. These results are comparable to the findings of Dowsett et al. (2016) (see Figure 2).

Discussion

Viewing multiple different images of a specific other-race face helped improve recognition of that face, with no transfer across identities. These results are consistent with findings regarding exposure to within-person variability in own-race faces (Andrews et al., 2015; Bindemann & Sandford, 2011, Dowsett et al., 2016, Menon et al., 2015b; Ritchie & Burton, 2017; White, Burton et al., 2014). This suggests that exposure to within-person variability might be key to facilitating familiarization with, and thus accurate recognition of, both own and other-race faces.

However, this method has a key limitation; as in Dowsett et al. (2016), participants knew that the target was always present in the lineup, precluding measurement of false alarms (see Megreya & Bindemann, 2015). The ability to accurately recognize that two photos belong to the same person is at least partially independent of the ability to recognize that two photos belong to different people; viewing multiple images improves matching on trials in which the sample and target images depict the same person but not on trials in which they depict different people (Ritchie & Burton, 2017; White, Burton et al., 2014). Further, Megreya and Burton (2007) found no correlation between accuracy on match and mismatch trials, both in a 1-to-1 and a 1-in-10 matching task.

Target-absent trials have external validity. In applied settings, perceivers (e.g., cashiers, border officials, eyewitnesses) cannot be sure whether the target identity (e.g., the culprit/person depicted in photo ID) is present. For example, when constructing a lineup for eyewitness identification, the police select a single suspect to be placed among known innocent distractors (see Kemp & White, 2016; Wells et al., 2000). Whether the suspect is the person previously seen by the witness is unknown. Ronald Cotton, a young Black man, was placed in such a lineup after being suspected of raping a Caucasian woman (Innocence Project, 2016). He was innocent, but when the victim identified Cotton as the culprit, the police unknowingly accepted this cross-race identification. To the extent that multiple images foster the formation of a robust abstract representation of identity, they might also help observers exclude images of a different person.

Alternatively, increased tolerance of variability might increase the number of false alarms when the target is absent. In Experiment 2 I included both target-present and target-absent trials to investigate whether within-person variability also helps participants recognize that the target identity is absent.

Experiment 2: Method

Participants

Thirty undergraduate students (18 female, M_{age}= 19.57, Range= 18-27) participated; this sample size allowed us to counterbalance the use of stimuli across trial types. Power analyses determined that with this sample size and alpha set to 0.05, the power to detect an effect comparable to that of Experiment 1 was 0.99 for the linear contrast analyses and 0.67 for the supplemental paired sample t-tests (2-tailed). As in Experiment 1, all participants reported minimal interaction; they reported having fewer than two Black friends, with 15 reporting no Black friends. An additional participant was excluded because of experimenter error. All participants provided written informed consent and received research credit for their participation.

Stimuli and Procedure

The same 30 identities were used as stimuli but each test array included only 28 of these identities so that each participant could be tested with two target-present and two target-absent arrays (images of the target-absent identities were removed from the array). Although the array remained the same across trials for each participant, it varied across participants because I rotated identities such that each identity served as a target for four different participants—twice as a target-present (P) identity and twice as a target-absent (A) identity. Participants completed these trials in one of two different orders (P-A-A-P or A-P-P-A) and were asked to identify whether the target was present in the lineup and, if so, to choose the correct image of the target. No feedback was given at any time.

To allow participants to complete the study in a timely fashion they were gradually presented with four, rather than six, different images of each target identity. Supplemental analyses of Experiment 1 based on four images revealed the same pattern of significant effects for both the linear contrasts and the paired t-tests as that reported for six images. Furthermore, paired t-tests were conducted comparing performance after four and six images for each identity. Not one of these tests was significant, all *p*s > .08.

Results

Target-Present Arrays

Accuracy. Figure 3 shows weak learning on target-present trials. Linear contrast analyses revealed that accuracy increased as more images were added of the second, $F(1,29) = 6.43$, $p = .017$, $\eta_p^2 = .18$, but not the first target, $F(1,29) = 0.11$, $p = .918$, $\eta_p^2 = .00$. Paired sample t-tests (2-tailed) confirmed that participants' performance was significantly better when they were able view four compared to only one image for the second target, $t(29) = 3.28$, $p = .003$ $d = 0.55$, but not the first, $p = .745$, perhaps because of such low performance when only one image of the second target was available (46% correct). On average, participants selected about two different identities across the four trials for each identity.

Misses vs. Misidentifications. Whereas only one type of error was possible in Experiment 1 (Misidentifications), on the target-present arrays of Experiment 2, there were two different types of errors that participants could make on each of the four trials: Misses and Misidentifications. Misses occur when participants incorrectly claim that the target is missing from the lineup; misidentifications occur when participants select the wrong person from the lineup, mistaking them to be the target identity. These have different implications in the real world (e.g., claiming that the bearer of ID does not match their photo and denying them border entry or age-restricted goods vs. wrongly accepting an invalid photo); therefore it is important to examine which type of mistake participants are making in this task. A 2 (Error Type: misses vs. misidentifications) x 2 (Identity: first vs. second) repeated measures ANOVA revealed a main effect of error type. Participants made more misidentifications ($M = 1.17$) than misses ($M = 0.40$), $F(1,29) = 11.27$, $p = .002$, $\eta_p^2 = .28$. Neither the main effect of identity nor the interaction between identity and error type were significant, *p*s > .07, indicating that for both the first and second target-present trials the most common error was misidentifications. To examine whether the number of misidentifications vs. misses differed as additional images were added, a chi-square analysis was conducted. The results of the chi-square indicated the type of errors made did not differ as a function of the number of images presented, χ^2 (3, $N = 94$) = 3.57, $p = .311$.

Target-Absent Arrays

Accuracy. The percentage of participants who correctly rejected the target from the lineup did not increase linearly as additional images were presented, *p*s > .27. Paired sample t-tests (2-tailed) confirmed that performance did not improve after viewing four images compared to just one, *p*s > .16. This is not a limitation of sample size; accuracy actually tended to decrease across trials, such that participants tended to make more false alarms as more images became available (see Figure 3). On each trial, between 57% and 73% of participants made a false alarm. On average, participants selected about two

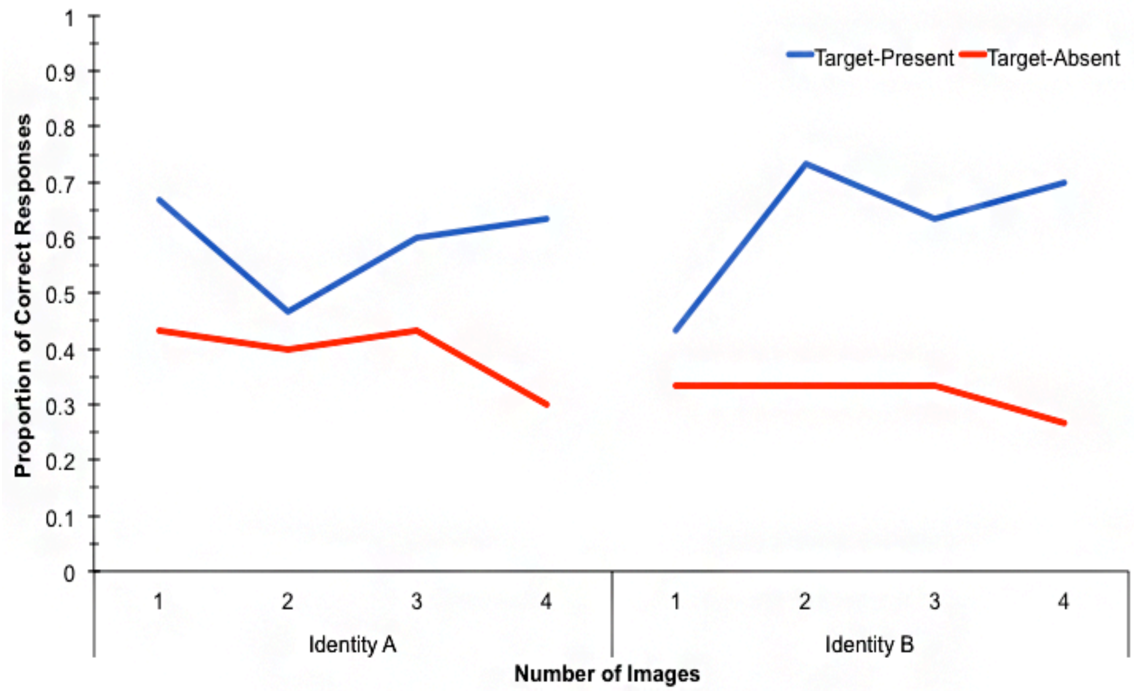

Figure 3. Proportion of participants who correctly selected the target on target-present (blue lines) and target-absent (red lines) trials.

different identities across the four trials for each identity. In summary, I found only weak evidence of improvement on target-present trials, but absolutely no evidence of improvement on target-absent trials.

Discussion

In Experiment 1, I found that viewing multiple different images of a specific other-race face helped improve recognition of that face. In that study the target was always known to be present. To resolve this limitation and increase external validity, in Experiment 2 I included both target-present and target-absent trials to investigate whether within-person variability also helps participants recognize that the target is absent. Here, I found weak evidence of learning on target-present trials; I observed significant improvement as additional images were added for the second identity with which participants performed the task, but not the first. This cannot be attributed to the difficulty of any particular target identity, as 30 different target identities were used across participants. On target-absent trials, I found absolutely no improvement after viewing multiple images of the target. This suggests that the results of Experiment 1 are specific to a situation in which the target is known to be present. However, when participants are uncertain whether the target is present, the extent to which exposure to multiple images improves recognition of the target decreases. Further, exposure to multiple images does not seem to improve the ability to detect when the target is missing from the lineup.

General Discussion

Forensic identity verification often involves observers and perceivers of different races. An abundance of research suggests that human perceivers have extreme difficulty recognizing other-race faces (Meissner & Brigham, 2001) and this is evident in the large number of wrongful conviction cases in which cross-race identification was a contributing factor (Innocence Project, 2016). Whereas recognizing an identity across a set of ambient images is especially challenging for unfamiliar other-race identities (Laurence et al., 2016), once we are familiar with an other-race face, we can easily recognize that identity despite changes in appearance. For example, Chinese adults sort images of known Caucasian celebrities with ease (Zhou & Mondloch, 2016). Understanding how a newly encountered other-race face becomes familiar is crucial to improving face matching in forensic settings and provides novel insights about models of face recognition. Exposure to within-person variability has been suggested to be a route to familiarization with own-race faces (Andrews et al., 2015; Bindemann & Sandford, 2011; Dowsett et al., 2016; Menon et al., 2015b; Ritchie & Burton, 2017). Therefore, the goal of the current studies was to investigate within-person variability as a viable route to familiarization with other-race faces.

In Experiment 1, I provide the first evidence that exposure to within-person variability aids in identity matching of other-race faces, at least when the observer is confident that the target is present. The proportion of participants who identified the target from the lineup increased as additional images were provided. Further, I found no evidence of transfer of learning across identities, suggesting that the improvement from exposure to within-person variability in appearance is identity-specific. This pattern of results is comparable to that previously observed for own-race faces (Dowsett et al., 2016). In fact, Figure 2 shows that my results mirror those of Dowsett et al. (2016) quite closely. This suggests that the processes by which a newly encountered face becomes familiar may be qualitatively the same for own-and other-race faces.

Such learning with own-race faces is thought to reflect the emergence of an abstract representation of a face (a Face Recognition Unit; FRU) (Bruce & Young, 1986), which is resistant to interference from variability in appearance attributable to environmental (e.g., lighting), camera (e.g., lens) and intrinsic (e.g., make-up, expression) factors (Burton, Jenkins, Hancock, & White, 2005). With each new encounter, tolerance for variability increases. My results suggest a similar process occurs for other-race faces.

In Experiment 2, I included target-absent trials to investigate whether within-person variability also helps to detect that the target identity is absent from the lineup. The results of Experiment 2 suggest that the learning observed in Experiment 1 is somewhat illusory. Under conditions of uncertainty, I found only a weak benefit of viewing multiple images when the target is present and no benefit when the target is absent. In Experiment 1, the only type of error participants could have made was a misidentification. However in Experiment 2, a new type of error was introduced on target-present trials: misses. Although less frequent than misidentifications, the introduction of misses at least partially accounts for the reduced improvement relative to Experiment 1, in which participants selected a face on every trial. Most notably, I found no evidence that adding exemplars helped participants to detect that the target identity was absent. Indeed, I observed a non-significant trend in the opposite direction, such that the proportion of participants' selecting an identity from the lineup despite the target being absent rose from about 60% to 70% after participants were presented with four images. Although non-significant results from the paired sample t-tests could be attributable to low power (0.67), the results of these supplemental analyses were consistent with corresponding linear contrast analyses, which had ample power (0.99) to find an effect comparable to Experiment 1. These findings are consistent with participants adopting a more liberal criterion in a delayed match-to-sample task after viewing two images attributed to the same identity than after viewing the same two images attributed to different identities (Menon, White & Kemp, 2015a), an effect that is enhanced by increased between-image variability (Menon et al., 2015b). Thus, while exposure to multiple images increases tolerance to variability in appearance, it allows participants to erroneously include images of non-target identities in their representation of the target (i.e., they become more willing to incorrectly accept between-identity differences as within-person variability).

Increasing tolerance to variability in appearance can be conceptualized as expanding attractor fields in multidimensional face space. According to Valentine (1991), vectors in multidimensional face space represent dimensions on which faces can vary (i.e., distance between the eyes, size of nose); the value of each face on these dimensions determines its location in face space. Thus, each identity we encounter is represented as a unique point in multidimensional face space. Expanding on this theory, Tanaka, Giles, Kremen and Simon (1998) proposed that each face is represented by an attractor field. These attractor fields represent the range of variability in appearance that will be perceived as belonging to that identity. Tanaka et al. (1998) emphasized the influence of distinctiveness on the size of a face's attractor field and Laurence et al. (2016) suggest that attractor fields for unfamiliar faces are influenced by face race. Both of these studies suggest that the size of a face's attractor field is influenced by the density of face space with faces in less densely clustered regions (i.e., distinctive faces, own-race faces) having larger attractor fields than faces in more densely clustered regions (i.e., typical faces, other-race faces). Accurate recognition of familiar faces despite within-person variability in appearance and for both own- and other-race faces can be conceptualized as familiar faces having large attractor fields; we are able to recognize a familiar person despite a wide range of within-person variability in their appearance. This means that when we encounter a familiar face in a new context (e.g., with a new hair colour), we can tolerate the variability in appearance because the attractor field is large enough for this new exemplar to fall within its boundaries.

The results of Experiment 1 and of Dowsett et al. (2016) can be conceptualized as expanding attractor fields for other-race and own-race faces, respectively; as participants were exposed to more images of the target, they were better able to tolerate the variability in the target's appearance and correctly select them from the lineup. However, two results of Experiment 2 suggest that the benefit of expanded attractor fields might be limited in a situation in which participants do not know whether the target was present. First, I found evidence of only weak learning on target-present trials. Second, and most notably, when the target was absent, exposure to multiple images made participants more tolerant to variability. This tolerance appears to have allowed an exemplar of a different identity to be brought into the attractor field, thus the number of false alarms did not decrease as images were added.

These studies have several strengths that support my conclusions. This is the first study to examine whether exposure to within person variability improves recognition of other-race faces. Whereas Dowsett et al.'s (2016) paradigm only included trials in which the target was known to be present, in Experiment 2, I included both target-present and target absent arrays. This allows for the examination of false alarms, providing novel insights about the ability to detect when a target is missing from a lineup. Including both target-present and target absent arrays also increases external validity; in applied settings, perceivers (e.g., cashiers, border officials, eyewitnesses) cannot be sure whether the target identity (e.g., the culprit/person depicted in photo ID) is present. Therefore, Experiment 2 tries to more closely mimic the task faced by those in forensic verification settings.

Additionally, Dowsett et al. (2016) used the same three target identities across all participants. I corrected this limitation by rotating through 30 different target identities. This ensures that my effects cannot be attributable to particular models being more or less difficult to recognize, increasing generalizability of the results. I also randomized the order in which participants viewed the training images. This makes it unlikely that improved performance reflects the utility of any one image (e.g., the one most similar to target image). Finally, all identities served as targets on both target-present and target-absent trials, eliminating any influence of identity on the effect of trial type. Collectively, these strengths give me confidence in my results.

Future Research

The goal of the current research was to replicate Dowsett et al's (2016) method to investigate whether exposure to multiple images improves the ability to select an other-race identity from a lineup. To tackle a limitation of their method, in Experiment 2, I included both target-present and target-absent to investigate whether exposure to within-person variability also improves detection that the target is missing from the lineup. The purpose of this research was not to examine the other-race effect and thus including own-race faces was beyond the scope of this study. Future research should investigate this question with own-race faces; it is unknown whether exposure to multiple images of a target would help participants detect when an own-race face is absent from the lineup. However, research with own-race faces has found that participants adopt a more liberal criterion after exposure to within-person variability (Menon et al., 2015a, Menon et al., 2015b). Thus, as with other-race faces, we would expect no improvement for own-race faces on target-absent trials.

Both Dowsett et al. (2016) and the current studies used lineups of 28 or 30 different identities. Such large lineups likely tap our ability to search for an identity in a crowd. However, a lineup of this size is not typically encountered in forensic settings. Although researchers have made some suggestions to improve lineup construction, there have not been specific recommendations made regarding the number of identities to include in a lineup (Wells et al., 1998), and therefore the number of lineups members actually used in forensic settings may vary. However, the majority of research examining eyewitness identification has used significantly less than 30 lineup members, often including only 6 (e.g., Wells, Steblay & Dysart, 2015). Identifying a target from a lineup of 30 images (a protocol that mimics finding a face in a crowd) may require a qualitatively different strategy than identifying a target from a lineup of only 6 identities. Therefore, future research should examine whether exposure to variability improves recognition in lineup sizes that more closely match those encountered in forensic settings.

The findings of Experiment 1 suggest that in a context in which the target is known to be present, exposure to natural variability in appearance seems to improve recognition of other-race faces. As can be seen in Figure 2, the pattern of improvement maps on quite perfectly to that found in Dowsett et al. (2016) with own-race faces. This suggests that the acquisition of familiarity for own-and other-race faces might share the same underlying processes. However, the question still remains as to whether we can learn other-race faces as efficiently as own-race faces. Future research should manipulate the duration of exposure and the amount of within-person variability that is necessary to learn an other-race compared to an own-race face.

Implications for Forensic Settings

Given the serious consequences of errors in identity matching (i.e., eyewitness misidentification, security risks etc.), understanding the mechanisms underlying such errors and designing protocols to reduce them is essential. In the context of eyewitness identification, interventions have focused mainly on improving the construction of lineups; researchers have advocated for the use of double-blind, sequentially (rather than simultaneously) presented lineups in which the suspect is placed among distractors of similar physical description that differ in appearance (see Wells et al., 2000 for review). In the context of border security, improvements in fraudulent document detection have forced fraudsters to switch from forging photo identification to using another person's identity, rendering the ability to match a face to a photograph even more important (ABC TV Science, 2015). Super-recognizers, those with an extraordinary ability to matching faces regardless of familiarity, have recently been recruited by border security and police departments to help combat this problem (ABC TV Science, 2015; Robertson, Noyes, Dowsett, Jenkins, & Burton, 2016). However, super-recognizers are a limited resource, and therefore finding an intervention to ease the difficulty of unfamiliar face matching for all forensic professionals, as well as eyewitnesses, is imperative. White, Burton et al., (2014) suggested that increasing the number of images on photo-ID might facilitate improved matching for own-race faces; participants were more accurate at matching when given four images of an identity rather than only one.

My results extend this possibility to other-race faces, but only when the observer knows the target is present, a rare circumstance in forensic settings. Indeed, viewing multiple images of an other-race face might actually increase effectiveness of false documents; providing evidence of one's variability in appearance (across identity cards and/or ambient images) might erroneously increase the perceiver's tolerance of images belonging to a similar-looking person. Likewise, whereas Experiment 1 suggests that the benefit of sequential (compared to simultaneous) photo arrays, in which eyewitnesses make yes/no judgements for each individual photo (Kemp & White, 2016; Wells et al., 2000), might be further enhanced by representing each identity with multiple images, Experiment 2 raises caution. Such a practice might also increase false identifications if the culprit is indeed absent.

References

ABC TV Science (Producer). (2015, February 24). *Catalyst: Episode 4* [Video Podcast]. Retrieved from http://www.abc.net.au/catalyst/vodcast/2015.htm

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *The Quarterly Journal of Experimental Psychology*, *68*, 2041-2050. doi:10.1080/17470218.2014.1003949

Bindemann, M., & Sandford, A. (2011). Me, myself, and I: Different recognition rates for three photo-IDs of the same person. *Perception*, *40*, 625-627. doi:10.1068/p7008

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology*, *73*, 105-116. doi:10.1111/j.2044-8295.1982.tb01795.x

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P. J., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, *5*, 339-360. doi:10.1037/1076-898X.5.4.339

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, *7*, 207-218. doi:10.1037/1076-898X.7.3.207

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327. doi:10.1111/j.2044-8295.1986.tb02199.x

Burton, A. M. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *The Quarterly Journal of Experimental Psychology*, *66*, 1467-1485. doi:10.1080/17470218.2013.800125

Burton, A. M., & Jenkins, R. (2011). Unfamiliar face perception. In Calder, A. J., Rhodes, G., Johnson, M. H., & Haxby, J. V. (Eds), *The Oxford Handbook of Face Perception* (pp.287-306). New York: Oxford University Press.

Burton, A. M., Jenkins, R., Hancock, P. J., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*, *51*, 256-284. doi:10.1016/j.cogpsych.2005.06.003

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods*, *42*(1), 286-291. doi:10.3758/BRM.42.1.286

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*, *10*, 243-248. doi:10.1111/1467-9280.00144

Dowsett, A. J., Sandford, A., & Burton, A. M. (2016). Face learning with multiple images leads to fast acquisition of familiarity for specific individuals. *The Quarterly Journal of Experimental Psychology*, *69*, 1-10. doi:10.1080/17470218.2015.1017513

Evans, J. R., Marcon, J. L., & Meissner, C. A. (2009). Cross-racial lineup identification: Assessing the potential benefits of context reinstatement. *Psychology, Crime & Law, 15*, 19–28. doi:10.1080/10683160802047030

Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. Nature Neuroscience, *4*, 845–850. doi:10.1038/90565

Gonzalez, C., & Madhavan, P. (2011). Diversity during training enhances detection of novel stimuli. *Journal of Cognitive Psychology*, *23*, 342-350. doi:10.1080/20445911.2011.507187

Hancock, P. J., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*, 330-337. doi: 10.1016/S1364-6613(00)01519-9

Hayward, W. G., Favelle, S. K., Oxner, M., Chu, M. H., & Lam, S. M. (2017). The other-race effect in face learning: Using naturalistic images to investigate face ethnicity effects in a learning paradigm. *The Quarterly Journal of Experimental Psychology*, *70*, 890-896. doi:10.1080/17470218.2016.1146781

Hole, G. J., George, P. A., Eaves, K., & Rasek, A. (2002). Effects of geometric distortions on face-recognition performance. *Perception*, *31*, 1221-1240.

Hussain, Z., Bennett, P. J., & Sekuler, A. B. (2012). Versatile perceptual learning of textures after variable exposures. *Vision Research*, *61*, 89-94. doi:10.1016/j.visres.2012.01.005

Innocence Project (2016). Retrieved from: http://www.innocenceproject.org/reevaluating-lineups-why-witnesses-make-mistakes-and-how-to-reduce-the-chance-of-a-misidentification/.

Jackiw, L. B., Arbuthnott, K. D., Pfeifer, J. E., Marcon, J. L., & Meissner, C. A. (2008). Examining the cross-race effect in lineup identification using Caucasian and First Nations samples. *Canadian Journal of Behavioural Science, 40*, 52–55. doi:10.1037/0008-400x.40.1.52

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313-323. doi:10.1016/j.cognition.2011.08.001

Kemp, R., & White, D. (2016). Face identification. In Groome, D., & Eysenck, M. W. (Eds.), *An introduction to Applied Cognitive Psychology*: Second Edition (pp. 39-70). London; New York: Roultledge.

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology*, *11*(3), 211-222.

Laurence, S., Zhou, X., & Mondloch, C. J. (2016). The flip side of the other-race coin: They all look different to me. *British Journal of Psychology, 107, 374-388*. doi:10.1111/bjop.12147

MacLin, O. H., & Malpass, R. S. (2001). Racial categorization of faces: The ambiguous race face effect. *Psychology, Public Policy, and Law*, *7*, 98–118. doi:10.1037/1076-8971.7.1.98

McGugin, R. W., Tanaka, J. W., Lebrecht, S., Tarr, M. J., & Gauthier, I. (2011). Race-specific perceptual discrimination improvement following short individuation training with faces. *Cognitive Science*, *35*, 330-347. doi:10.1111/j.1551-6709.2010.01148.x

Megreya, A. M., & Bindemann, M. (2015). Developmental improvement and age-related decline in unfamiliar face matching. *Perception*, *44*, 5-22. doi:1 0.1068/p7825

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*, 865-876. doi:10.3758/BF03193433

Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception & Psychophysics*, *69*, 1175-1184. doi:10.3758/BF03193954

Megreya, A. M., & Burton, A. M. (2008). Matching faces to photographs: poor performance in eyewitness memory (without the memory). *Journal of Experimental Psychology: Applied*, *14*, 364-372. doi:10.1037/a0013464

Megreya, A. M., Sandford, A., & Burton, A. (2013). Matching Face Images Taken on the Same Day or Months Apart: The Limitations of Photo ID. *Applied Cognitive Psychology*, *27*, 700-706. doi:10.1002/acp.2965

Megreya, A. M., White, D., & Burton, A. M. (2011). The other-race effect does not rely on memory: Evidence from a matching task. *The Quarterly Journal of Experimental Psychology*, *64*, 1473-1483. doi:10.1080/17470218.2011.575228

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

Meissner, C. A., Susa, K. J., & Ross, A. B. (2013). Can I see your passport please? Perceptual discrimination of own-and other-race faces. *Visual Cognition*, *21*, 1287-1305. doi:10.1080/13506285.2013.832451

Meissner, C. A., Tredoux, C. G., Parker, J. F., & MacLin, O. H. (2005). Eyewitness decisions in simultaneous and sequential lineups: A dual-process signal detection theory analysis. *Memory & Cognition, 33,* 783–792. doi:10.3758/BF03193074

Menon, N., White, D., & Kemp, R. I. (2015a). Identity-level representations affect unfamiliar face matching performance in sequential but not simultaneous tasks. *The Quarterly Journal of Experimental Psychology*, *68*, 1777-1793. doi: 10.1080/17470218.2014.990468

Menon, N., White, D., & Kemp, R. I. (2015b). Variation in Photos of the Same Face Drives Improvements in Identity Verification. *Perception*, *44*, 1332-1341. doi:10.1177/0301006615599902

Mondloch, C. J., Elms, N., Maurer, D., Rhodes, G., Hayward, W. G., Tanaka, J. W., & Zhou, G. (2010). Processes underlying the cross-race effect: An investigation of holistic, featural, and relational processing of own-race versus other-race faces. *Perception*, *39*, 1065-1085. doi: 10.1068/p6608

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *The Quarterly Journal of Experimental Psychology*, *70*(5), 897 905. doi:10.1080/17470218.2015.1136656

Robertson, D. J., Noyes, E., Dowsett, A. J., Jenkins, R., & Burton, A. M. (2016). Face Recognition by Metropolitan Police Super-Recognisers. *PLoS ONE*, *11*(2), 1-8. Doi: 10.1371/journal.pone.0150036

Tanaka, J. W., Giles, M., Kremen, S., & Simon, V. (1998). Mapping attractor fields in face space: The atypicality bias in face recognition. Cognition, *68*, 199–220. doi:10.1016/S0010-0277(98) 00048-1

Tanaka, J., & Pierce, L. (2009). The neural plasticity of other-race face recognition. *Cognitive Affective & Behavioral Neuroscience*, *9*, 122-131. doi:10.3758/CABN.9.1.122

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43*A, 161–204. doi:10.1080/14640749108400966

Watson, T., Robbins, R., & Best, C. (2014). Infant perceptual development for faces and spoken words: An integrated approach. *Developmental Psychobiology*, *56*, 1454-1481. doi:10.1002/dev.21243

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist*, *55(6)*, 581- 598.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, *22*(6), 1-37.

Wells, G. L., Steblay, N. K., & Dysart, J. E. (2015). Double-blind photo lineups using actual eyewitnesses: An experimental test of a sequential versus simultaneous lineup procedure. *Law and Human Behavior*, *39*, 1-14. doi: 10.1037/lhb0000096

White, D., Burton, A. M., Jenkins, R., & Kemp, R. I. (2014). Redesigning photo-ID to improve unfamiliar face matching performance. *Journal of Experimental Psychology: Applied*, *20*, 166-173. doi:10.1037/xap0000009

White, D., Dunn, J. D., Schmid, A. C., & Kemp, R. I. (2015). Error Rates in Users of Automatic Face Recognition Software. *PLoS ONE*, *10*, 1-14. doi:10.1371/journal.pone.0139827

White, D., Kemp, R. I., Jenkins, R., Matheson, M., & Burton, A. M. (2014). Passport Officers' Errors in Face Matching. *PLoS ONE*, *9*, 1-6. doi:10.1371/journal.pone.0103510

Wright, D. B., Boyd, C. E., & Tredoux, C. G. (2003). Inter-racial contact and the own-race bias for face recognition in South Africa and England. *Applied Cognitive Psychology, 17*, 365–373. doi:10.1002/acp.898

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology*, *102*, 959-974. doi: 10.1111/j.2044-8295.2011.02045.x

Young, A. W., & Burton, A. M. (in press). *Recognizing faces*. *Current Directions in Psychological Science*.

Zhou, X., & Mondloch, C. (2016). Recognizing 'Bella Swan' and 'Hermione Granger': No own-race advantage in recognizing photos of famous faces. *Perception, 45*, 1426-1429. doi:10.1111/bjop.12147

Appendix 1

Brock University
Research Ethics Office
Tel: 905-688-5550 ext. 3035
Email: reb@brocku.ca

Social Science Research Ethics Board

Certificate of Ethics Clearance for Human Participant Research

DATE:	October 24, 2016		
PRINCIPAL INVESTIGATOR:	MONDLOCH, Cathy - Psychology		
FILE:	15-232 - MONDLOCH		
TYPE:	Ph. D.	STUDENT: SUPERVISOR:	Xiaomei Zhou Cathy Mondloch

TITLE: Recognizing Faces in Photographs

ETHICS CLEARANCE GRANTED

Type of Clearance: MODIFICATION   Expiry Date: 3/31/2017

The Brock University Social Sciences Research Ethics Board has reviewed the above named research proposal and considers the procedures, as described by the applicant, to conform to the University's ethical standards and the Tri-Council Policy Statement.

Modification: Computer stimulus and task.

The Tri-Council Policy Statement requires that ongoing research be monitored by, at a minimum, an annual report. Should your project extend beyond the expiry date, you are required to submit a Renewal form before **3/31/2017**. Continued clearance is contingent on timely submission of reports.

To comply with the Tri-Council Policy Statement, you must also submit a final report upon completion of your project. All report forms can be found on the Research Ethics web page at http://www.brocku.ca/research/policies-and-forms/research-forms.

In addition, throughout your research, you must report promptly to the REB:

a) Changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) All adverse and/or unanticipated experiences or events that may have real or potential unfavourable implications for participants;
c) New information that may adversely affect the safety of the participants or the conduct of the study;
d) Any changes in your source of funding or new funding to a previously unfunded project.

We wish you success with your research.

Approved:

Jan Frijters, Chair
Social Science Research Ethics Board

Note: Brock University is accountable for the research carried out in its own jurisdiction or under its auspices and may refuse certain research even though the REB has found it ethically acceptable.

If research participants are in the care of a health facility, at a school, or other institution or community organization, it is the responsibility of the Principal Investigator to ensure that the ethical guidelines and clearance of those facilities or institutions are obtained and filed with the REB prior to the initiation of research at that site.

Appendix 2

Participant #______

QUESTIONNAIRE: RACE

DOT ________________

Thank you for participating in our research. We recognize that individuals may differ in their ability to recognize faces. Some of these differences may be attributable to how much we experience different kinds of faces on a daily basis. Please take a few moments to complete the following questionnaire.

Your responses will be confidential.

PERSONAL INFORMATION

1. **Date of birth**: ……………………………..……………………………….
2. **Mark your ethnic group:**

   - ❑ Caucasian
   - ❑ Asian
   - ❑ Aboriginal
   - ❑ African Canadian
   - ❑ Other (Please specify) __________________________

3. How many people live in your household (including yourself):

   a) Home _________
   b) University __________

4. Please indicate how many of those people (both at home and at University) are in each of the following groups:

   a) Caucasian _________
   b) African Canadian __________

5. Think about family members with whom you have regular contact (at least once per month). How many of those people are in each of the following ethnic groups:

   a) Caucasian _________
   b) African Canadian __________

6. Please estimate how many hours you usually spend per week interacting with people in each of the following groups:

a) Caucasian _________
b) African Canadian __________

7. With how many African Canadian adults do you have contact in a typical week?
____

8. In your opinion, in the last 5 years, how much experience have you had with African Canadian individuals?

1	2	3	4	5
minimal	some	moderate	a lot	extensive

9. Think of up to 10 friends with whom you spend the most time. Of these 10 friends:

How many are Caucasian? _______
How many are African Canadian? _______
How many are any other race outside of Caucasian and African Canadian?
____

10. Please write down the FIRST NAME (only) of up to 10 African Canadian friends:

____

____

____

11. Please provide any additional information that would indicate extensive experience with African Canadian individuals.

Appendix 3

Participant Number ___________________
Ethnicity (e.g. White, Black, Asian): __________________________________

Were any of the faces in the experiment familiar? Please circle

Yes / No

If yes, please indicate the name/s* of the individual/s that you recognized:

____

____

*If you can't recall a name then please write down other information related to that person.

____

____