Are They Right or Wrong? Investigating the Ability to Judge the Accuracy of Eyewitnesses in Same- and Other-Race Identifications

by

Katherine E. L. Riess

A Thesis Submitted in Partial Fulfillment

of the Requirements of the Degree of

Masters of Arts

in

The Faculty of Social Science and Humanities

Criminology Program

University of Ontario Institute of Technology

August 2012

Certificate of Approval

Submitted by **Katherine Riess**

In partial fulfillment of the requirements for the degree of

Master of Arts in **Criminology**

Date of Defence: **2012/08/08**

Thesis title: Are They Right or Wrong? Investigating the Ability to Judge the Accuracy of Eyewitnesses in Same- and Other-Race Identifications

The undersigned certify that the student has presented her thesis, that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate through an oral examination. They recommended this thesis to the Office of Graduate Studies for acceptance.

Examining Committee

Dr. Barbara Perry	Chair of Examining Committee
Dr. Ron Hinch	External Examiner
Dr. Brian Cutler	Research Supervisor
Dr. Amy Leach	Examining Committee Member

As research supervisor for the above student, I certify that I have read and approved changes required by the final examiners and recommend the thesis for acceptance:

_____

Dr. Brian Cutler
Research Supervisor

Abstract

Past research has investigated the cross-race effect in the context of eyewitnesses and jury decision-making. The main goal of my thesis was to gain further insight into participants' knowledge of the cross-race effect and how this impacted participants' discrimination of same- and cross-race identifications. One hundred fifty-nine undergraduate students from UOIT viewed a series of showup identification videos. I found that participants were better able to discriminate accurate from inaccurate same-race identifications than cross-race identifications. However, participants believed White witnesses more and found them more credible than South Asian witnesses. Further research should investigate other conditions that influence people's abilities to discriminate accurate from inaccurate eyewitness identifications.

Key Words: Jury decision making, eyewitnesses, cross-race effect, juror knowledge.

Acknowledgements

I would like to thank my supervisor, *Dr. Brian Cutler*, for his never-ending guidance, support, and patience throughout my Masters degree. Dr. Cutler, you guided me through the past two years of my academic career with a steady and nurturing hand, showing me that I have the ability to accomplish anything I set my mind to. Without your guidance, this research would not be what it is today. I hope I can embody the professionalism, respect, and dignity you showed me in all aspects of my life. I am so honoured to graduate before you.

To *Dr. Amy-May Leach*, thank you for your encouragement and wisdom on this project. I appreciate all the input and assistance you provided, helping me to look deeper at my research and take my ideas even further. It was an honour to work with such an amazing researcher and person.

To my *research assistants*, thank you for your hard work and dedication to this research. The hours you spent in the lab helped make this research possible, and I value the lessons I learned from you over the last two years.

To my *family and friends*, I appreciate the support you have given me over the last two years more than you could know. *Mom and Dad*, you give me strength and wisdom that I cannot learn from a textbook. To my fiancé, *Kevin*, you listened to me go on and on about my research and you always believed in me.

Thank you.

Table of Contents

List of Tables

List of Figures

List of Appendices

Are They Right or Wrong? Investigating the Ability to Judge the Accuracy of Eyewitnesses and Juror Sensitivity to the Cross-Race Effect

Eyewitnesses play an important role in the criminal justice system, yet their identifications are frequently erroneous (e.g., Steblay Dysart, Fulero, & Lindsay, 2001; Steblay, Dysart, Fulero, & Lindsay, 2003; Yarmey, Yarmey, & Yarmey, 1996). In addition, research has shown that jurors do not understand the factors affecting eyewitness memory and eyewitness identification. Even when they do understand how some factors influence eyewitness identification, jurors do not take those factors into account when judging the accuracy and credibility of eyewitnesses (Cutler, Penrod, & Dexter, 1990; Desmarais & Read, 2011; Martire & Kemp, 2009). This is an important issue, as judges and jurors routinely evaluate eyewitness identifications, including those made in showup situations (Steblay et al., 2003). The purpose of my thesis was to examine participants' abilities to discriminate between accurate and inaccurate eyewitnesses and to investigate whether participants were sensitive to the cross-race effect. In order to provide a context for my study, below I review existing literature surrounding eyewitness identifications from showups, the accuracy of jurors' judgements of eyewitnesses, and jurors' sensitivity to cross-race effect as it pertains to witness identifications.

The Importance of Eyewitness Identifications

In 1971, five Black men from Florida were charged with the murder of Khomas Revels (Meissner & Brigham, 2001). Five White eyewitnesses implicated those men (i.e., the Quincy Five). The State argued that this was proof beyond a reasonable doubt and that there was no better evidence than five corroborated statements from

"unprejudiced witnesses" (Meissner & Brigham, 2001, p. 3). Two of these five men were wrongfully convicted for the murder of Khomas Revels, despite the lack of physical evidence tying them to the crime. This case is just one of the many wrongful convictions in North America. According to the literature, 75% of wrongful convictions in the United States that are rectified by exonerations are due to erroneous eyewitness identifications (deLone, 2011; Huff, 2004). Thus, not only are eyewitnesses making errors in their identifications, but judges and jurors believe these errors.

Cross-Race Recognition

Eyewitnesses are more accurate when identifying members of their own races than members of different races (Meissner & Brigham, 2001). This is known as the cross-race effect or the own-race bias. In reviewing this literature, Meissner and Brigham (2001) meta-analyzed a total of 39 published and unpublished articles. The majority of samples used for this analysis consisted of White and Black witnesses and suspects, although some studies used other races. Ninety-one percent of the studies made use of a recognition paradigm (i.e., participants viewed faces and later recognized them from a larger set of faces) and 9% used lineup identification procedures. Meissner and Brigham (2001) concluded that, generally, own-race identification attempts produced more correct identifications and fewer false identifications than did cross-race identification attempts. Meissner and Brigham (2001) concluded that the cross-race effect is a generalizable and reliable phenomenon across a variety of races/ethnicities, cultures, and situations.

Evaluations of Same- and Cross-Race Eyewitness Identifications

A meta-analysis by Desmarais and Read (2011) examined the beliefs of respondents regarding eyewitnessing factors. It is important to examine laypersons'

knowledge because it is typically laypersons (i.e., jurors) who evaluate eyewitnesses in the criminal justice system. Thus, the researchers examined 23 different surveys, from 15 different studies, that used direct methods to assess the knowledge of laypersons. Samples were drawn from the United States, Canada, the United Kingdom, and Australia. Desmarais and Read found that there was a general understanding that cross-race identifications were less accurate than same-race identifications.

Many laypeople are knowledgeable that own-race identifications are more likely to be accurate and more credible than other-race identifications (Desmarais & Read, 2011; Read & Desmarais, 2009). Awareness, however, does not necessarily translate into sensitivity when jurors are evaluating identifications during a trial. Abshire and Bornstein (2003) investigated the cross-race effect with Black and White witnesses and mock jurors. Through stimulus photos assessed by the mock-jurors, the researchers were able to manipulate the races of the eyewitnesses (either Black or White). The race of the defendant was held constant (Black). Black and White undergraduate students, who acted as jurors in this study, listened to an audio-taped murder trial. Photos of each of the key 'players' in the trial (i.e., witnesses, defendant, lawyers, police detective) were shown to half of the participants, and the other half were kept unaware of the races of the trial participants. The authors also evaluated jurors' perceptions of race in the justice system and if jurors were aware that cross-race identifications were less accurate than same-race identifications. Abshire and Bornstein (2003) found that jurors did not take the cross-race effect into account when judging the accuracy and credibility of eyewitnesses, even if they were aware of its impact. Overall, White jurors rendered more guilty verdicts than did Black jurors. Black jurors rated Black witnesses as more credible than did White

witnesses; White jurors, however, did not rate White witnesses as more credible than Black witnesses.

Although previous research has not examined how the cross-race effect influences people's abilities to discriminate accurate from inaccurate eyewitnesses (referred to below as "accuracy judgments"), there has been previous research on accuracy judgments of eyewitnesses in general. Martire and Kemp (2009) investigated how expert testimony influenced jurors' accuracy and knowledge of factors affecting eyewitness identifications. To do this, they employed a two-stage research design. In the first phase, participants watched videotapes of crimes and attempted lineup identifications. In the second phase, participants watched mock trials that included eyewitness identifications from the first phase. The mock trials included a judicial instruction video, congruent or incongruent expert testimony, or no instructions. Overall, jurors performed significantly better than chance when evaluating the accuracy and inaccuracy of eyewitnesses. The instruction condition (control, congruent expert, incongruent expert, and judicial instruction) did not impact this ability. The purpose of the above research was to evaluate the impact of eyewitness expert testimony and judicial instructions on juror judgement accuracy (Martire & Kemp, 2009). Focusing on the control group (no instruction), the ethnic makeup of the witnesses, suspects, and participant-jurors was not discussed or evaluated. It is possible that this information could have impacted juror judgement accuracy. As cross-race identifications are more difficult to make than same-race identifications (Meissner & Brigham, 2001), it is important to evaluate whether jurors are sensitive to this phenomenon and whether the races of the witness and suspect impact juror judgement accuracy.

Overview of the Present Study

The purpose of the present research is twofold. First, I examined participants' abilities to discriminate between accurate and inaccurate eyewitnesses. Second, I examined whether participants were sensitive to the cross-race effect with self-identified White and South Asian witnesses and manipulated suspect races (White, South Asian, and No Race Information) and how the phenomenon impacted participants' accuracy judgements. In my study, I presented participants with 20 videos of showup identifications. Participants viewed the videos in a randomized order, judged the accuracy and credibility of each witness, and then rated their own confidence in their judgements. Once all of the videos had been viewed, participants were asked to rate how influential several factors were on their decisions to label the witnesses as correct or incorrect.

My study contributes uniquely to the literature in at least four ways. First, there is no research investigating participants' abilities to discriminate accurate from inaccurate eyewitness identifications from showups. Showups are eyewitness identification procedures in which a witness views just one person and is asked if this is the person they saw commit the crime (Steblay et al., 2003). The showup is different from a lineup in several ways. Because a lineup includes fillers (i.e., other persons or photos or other persons who are known to be innocent) and a showup does not, a lineup offers more protection for an innocent suspect (Steblay et al., 2003). Also, showups can be suggestive. Showups usually happen shortly after the crime and at or near the scene of the crime and can give the impression that the police think the suspect is the perpetrator, even if the suspect is innocent. In support of using showups, however, Steblay and

colleagues (2003) found that showups yielded a higher percentage of correct rejections than lineups. This indicates that, when faced with a showup and an innocent suspect, participants were more likely than when faced with a lineup to correctly say that the person before them was not the one they saw commit the crime.

Second, I examined participants' abilities to discriminate accurate from inaccurate same- and other-race identifications. I know of no published study that investigates this issue. Some researchers have investigated juror knowledge and sensitivity to cross-race identifications (Abshire & Bornstein, 2003; Desmarais & Read, 2011; Martire & Kemp, 2009). In my design I am able to investigate if the results of Abshire and Bornstein (2003), who examined juror sensitivity to the cross-race effect, extend to defendants of other races.

Third, most studies of cross-race identification have examined White and Black witnesses and targets (Meissner & Brigham, 2001). By contrast, my study used White and South Asian witnesses, and I manipulated perceived suspect race using a photo of a White or South Asian person. Use of different races can help determine the generalizability of the findings concerning same- versus cross-race identifications (Meissner & Brigham, 2001).

Fourth, I included measures of both perceived and actual accuracy. An important limitation of the study by Abshire and Bornstein (2003) is that it did not include a measure of judgement accuracy. In other words, the researchers could not examine whether the jurors were accurate or inaccurate in their assessments of the accuracy of the eyewitness. It is important to investigate both perceived and actual accuracy as they are both interesting psychological issues. Actual accuracy tells us about peoples' cognitive

abilities to evaluate eyewitnesses. Perceived accuracy tells us about what factors people believe influences eyewitness identification.

Hypotheses

1. Consistent with the results of Abshire and Bornstein (2003), I hypothesized that same-race showup identifications would be perceived as more accurate and more credible than other-race showup identifications (regardless of the actual accuracy of those perceptions).
2. Martire and Kemp (2009) found that participants performed better than chance when distinguishing between accurate and inaccurate eyewitnesses, but overall performance was weak (participants were on average 63.6% correct). The majority of actual accuracy studies, in contrast, have found that participants did not perform significantly better than chance (Beaudry, 2008; Lindsay, Wells, & O'Connor, 1989; Wells, Lindsay, & Ferguson, 1979). Therefore, consistent with the results of the majority of previous research (e.g., Wells, Lindsay, & Ferguson, 1979), I predicted that students would not discriminate accurate from inaccurate showup identifications beyond chance levels, regardless of whether identifications were same- or other-race.
3. Because Abshire and Bornstein (2003) found that participants were not sensitive to the cross-race effect, I expected that my participants would not be sensitive to cross-race effect. Therefore, I hypothesized that I would find nonsignificant differences between the average accuracy for evaluations of the accuracy of same-race and other-race identifications.

Method

Design

I used a 2 (Witness Race: White vs. South Asian) x 3 (Suspect Race: No information vs. White vs. South Asian) mixed factorial design. The within-subjects factor was Witness Race, and the between-subjects factor was Suspect Race.

Participants

One hundred fifty-nine participants aged 16 to 27 years (M = 18.92 years, SD = 1.65) were recruited from the University of Ontario Institute of Technology participant research pool. Participants were awarded 1% extra credit towards their specified Psychology course. The participants included men and women (71 males, 88 females). The ethnic origins of participants were Arab/West Indian (11.3%), Black (11.9%), Caucasian (37.1%), Chinese (5%), Filipino (0.6%), Hispanic (1.9%), South Asian (17.0%), South East Asian (4.4%), Other (8.8%), and Undisclosed (1.9%).

Materials

Eyewitness Identification Videos. The videos included in this study were short clips (M = 20.21 seconds, SD = 7.79) of actual showup identifications from a previous study (Smith, Leach, & Cutler, 2011). In Smith et al.'s study, participants were brought into the lab under the guise that they were writing an essay on an unrelated topic. While in the lab, participants witnessed a confederate take the researcher's bag. When the researcher returned to the room and noticed that his or her bag was missing, the researcher asked the participant what happened. The researcher then took the participant to another room where either the thief or an innocent person was sitting. The researcher asked the participant if this was the person who took the experimenter's bag and how

confident he or she was in his or her decision. Every participant's identification and confidence statement were videotaped using a hidden video camera, and consent was obtained to use these videos in future research.

The only portion of the videos used for this study was the eyewitnesses' identifications and confidence ratings. The participant-jurors could only see the experimenter and the eyewitness during the videos. The suspect was cropped out of the frame so that I could manipulate the race of the suspect. I used only videos in which the eyewitness made a positive identification (i.e., said that, yes, that is the person that took that bag). There were 10 correct identifications and 10 incorrect identifications, for a total of 20 videos. These 20 videos were balanced for witness race (eight white females, two white males, two South Asian females, and eight South Asian males). After conducting the present study, one South Asian witness video was removed because it was very unclear as to whether the witness had made an identification. After this removal, 10 correct identifications and 9 incorrect identifications remained. To create the appearance of a same- or other-race identification, the identifications were preceded by a photo of a suspect who was obviously White or South Asian (as defined by each person whose photo was used). In the control condition, no information regarding the race of the suspect was provided to participants.

Dependent variables. After viewing each video, participants were asked to decide if the witness was correct or incorrect. Each participant was then asked to rate his or her confidence in the decision on a scale from one (not at all confident) to five (very confident). Finally, each participant was asked to rate the credibility of each witness on a scale from one (not at all credible) to five (very credible; Appendix B).

After viewing all of the videos, participants were asked to rate the influence of various factors on their decisions to say that the witnesses were correct or incorrect. These factors were experimenter-related (i.e., experimenter's behaviour, what the experimenter did or did not say), witness-related (i.e., what the witness looked like, witness behaviour, witness confidence, how long it took the witness to make an identification, how carefully the witness looked at the suspect, how nervous the witness appeared, what the witness did or did not say, witness race), and suspect-related (i.e., suspect race). These factors were all rated on a five-point Likert scale, from one (the factor had no influence) to five (the factor had a very strong influence). Participants were also asked whether or not they agreed with the statement, "If the race of the suspect is different from the race of the witness, it is more difficult for the witness to make a correct identification" using a five-point Likert scale, from one (completely disagree) to five (completely agree; Appendix C).

Procedure

Upon entering the lab, participants, in groups of up to eight, were directed to sit at the cluster of tables in the middle of the main lab room. The researcher present (either myself or a Research Assistant) asked participants to read and sign informed consent forms if they agreed to participate (Appendix E). Once consent was obtained, participants were directed to sit at an individual desk with a computer (stationed around the parameter of the main lab room) and told how to use the laptops. The experimenter then launched the Media Lab program (the program used to run the study). I used a random numbers table generated on random.org to randomly assign participants to conditions.

The participants began the study by answering demographic questions in regards to age, gender, ethnicity, faculty, and year of program (Appendix F). The videos began once these questions were answered. Participants viewed a series of 20 videos, and each video was followed by three questions – whether participants thought the identification was correct or incorrect, how confident were they in their decision, and how credible they found each witness. The program used to play the videos for participants randomized the order of videos every time it was launched; therefore, each participant was shown the videos in random order.

After viewing all 20 videos and answering the corresponding questions, participants were asked to rate a series of factors on the level of influence each factor had on their judgements and credibility ratings. Once each factor was rated and participants answered the question regarding race identifications, participants were instructed by the program to notify the researcher that they were finished. The debriefing form was given to participants at this time (Appendix D). After reading the debriefing form, participants were asked if they had any questions or concerns. Upon exiting the lab, participants were awarded 1% extra credit toward their Psychology classes for participating in the study, and they were asked to keep the true nature of the study to themselves so as to not bias any participants that may participate in the study in the future. From start to finish, the study took participants between 30 minutes and 40 minutes to complete.

Results

Perceived Accuracy and Credibility of Same- and Other-Race Witnesses

I found considerable variability in accuracy judgements. Some participants perceived all witnesses to have made inaccurate identifications, whereas others perceived 90% of the witnesses to have made accurate identifications. On average, participants perceived witnesses to be accurate in their identifications roughly half of the time, regardless of whether they were same-or other-race identifications (M =.49, SD = .22).

I hypothesized that participants would perceive same-race identifications as more accurate than other-race identifications. I conducted a Suspect Race (No Information v. White v. South Asian) x Witness Race (White v. South Asian) mixed-factors ANOVA, with eyewitness race as the within-subjects factor. I found a significant main effect of witness race, indicating that participants perceived the White witnesses to be more accurate than the South Asian witnesses, $F(1,156) = 179.616$, $p < .001$. The main effect of suspect race, however, was not significant, $F(2,156) = 1.119$, $p = .329$. Contrary to my hypothesis, the interaction between witness race and suspect race was not significant, $F(2,156) = .602$, $p = .549$ (see Figure 1).

The results for witness credibility were very similar to those for perceived accuracy. I conducted the same mixed-factors ANOVA for credibility ratings. The main effect of witness race was significant, indicating that the White witnesses were rated as more credible than the South Asian witnesses, $F(1,156) = 262.266$, $p < .001$. The main effect of suspect race, like perceived accuracy, was not significant, $F(2,156) = 1.034$, $p = .358$. The interaction between witness race and suspect race was also not significant, $F(2,156) = .212$, $p = .810$ (see Figure 2).

To investigate if these results are qualified by participant race or gender, I conducted a Participant Race (White v. South Asian) x Participant Gender (Male v. Female) x Suspect Race (No Information v. White v. South Asian) x Witness Race (White v. South Asian) mixed-factors ANOVA with Witness Race as the within-subjects factor using perceived accuracy the dependent variable. I excluded data from participants who identified themselves as other than White or South Asian. Neither participant gender nor ethnicity interacted significantly with witness race or suspect race conditions, suggesting that the results are the same for White and South Asian participants.

Actual Accuracy Rates of Participant Judgements

The actual accuracy rates are summarized in Table 1. On average, participants performed relatively close to the level of chance, indicating that they had difficulty determining which witnesses made accurate identifications and which witnesses made inaccurate identifications. I hypothesized that participants would not discriminate accurate from inaccurate identifications beyond chance levels, regardless of whether the identification was same- or other-race. To test this hypothesis, I performed a series of one-way *t*-tests. Overall accuracy of participants was .50 (*SD* = .08). This accuracy rate did not differ significantly from chance (.50), $t(158) = .292$, $p = .771$. Accuracy for same-race identifications was .48 (*SD* = .15), which did not differ significantly from chance, $t(104) = -1.456$, $p = .148$. The accuracy rate for other-race identifications was .53 (*SD* = .13), which was significantly higher than chance, $t(104) = 2.174$, $p = .032$. Thus, support for my hypothesis was mixed. For exploratory purposes, I assessed whether accuracy rates differed from chance as a function of suspect race and witness race. With respect to suspect race, I found that evaluations of accuracy for White suspects, South

Asian suspects, and suspects for which no race information was given did not differ significantly from chance. For White suspects, the accuracy rate was .52 (*SD* = .09), *t*(54) = 1.266, *p* = .211. For South Asian suspects, the accuracy rate was .50 (*SD* = .08), *t*(49) = -.098, *p* = .922. For suspects for which no race information was given, the accuracy rate was .49 (*SD* = .06), *t*(53) = -1.124, *p* = .266. With respect to witness race, accuracy was significantly lower than chance when the witness was South Asian (*M* = .43, *SD* = .11), *t*(158) = -8.120, *p* < .001 but significantly better than chance when the witness was White (*M* = .57, *SD* = .12), *t*(158) = 7.368, *p* < .001.

I hypothesized that I would find nonsignificant differences between the average accuracy for evaluations of the accuracy of same-race and other-race identifications. To investigate this, I conducted a Suspect Race (No Information v. White v. South Asian) x Witness Race (White v. South Asian) mixed-factors ANOVA for Actual Accuracy rates, with Witness Race as the within-subjects factor. Participants were better able to distinguish between accurate and inaccurate White witnesses (*M* = .57, *SD* = .12) than South Asian witnesses (*M* = .43, *SD* = .11). The main effect of suspect race was not significant, *F*(2,156) = 1.780, *p* = .172. Contrary to my hypothesis, the interaction between witness race and suspect race was significant, *F*(2,156) = 5.793, *p* = .004 (see Figure 3). To understand the interaction, I conducted post hoc analyses. Participants' accuracy rates with White witnesses did not vary across suspect race conditions, *F*(2,158) = 1.305, *p* = .274. However, participant accuracy rates did differ significantly across suspect race conditions, *F*(2,158) = 7.021, *p* = .001. Post hoc tests indicate that the significant differences occurred between the No Information (*M* = .41, *SD* = .09) and White Suspect (*M* = .47, *SD* = .12) conditions, *p* = .012, and between the White Suspect

and South Asian Suspect (M = .40, SD = .12) conditions, p = .002, but significant differences did not occur between the No Information and South Asian Suspect conditions, p = .797. Accuracy rates were highest in the White suspect condition.

To investigate if the ANOVA results are qualified by participant race or gender, I conducted a Participant Race (White v. South Asian) x Participant Gender (Male v. Female) x Suspect Race (No Information v. White v. South Asian) x Witness Race (White v. South Asian) mixed-factors ANOVA with Witness Race as the within-subjects factor using perceived accuracy and actual accuracy as the dependent variable. I excluded data from participants who identified themselves as other than White or South Asian. Neither participant gender nor ethnicity interacted significantly with witness race or suspect race conditions, suggesting that the results are the same for White and South Asian participants.

Confidence

To investigate confidence ratings, I conducted a Suspect Race x Witness Race mixed-factors ANOVA. I found that the main effects of witness race and suspect race were not significant, $F(1,156) = 1.733$, p = .190, and $F(2,156) = 1.323$, p = .269, respectively. The interaction between witness race and suspect race was also not significant, $F(2,156) = .900$, p = .409 (see Figure 4).

Self-Reported Influences on Judgements

Participants rated witness confidence (M = 4.58, SD = .84), witness behaviour (M = 4.45, SD = .85), what the witness did or did not say (M = 4.21, SD = .87), and how long it took the witness to make an identification (M = 4.09, SD = .95) as the most influential factors on their judgements (see Table 2). Surprisingly, the factors rated as least

influential were what the witness looked like (M = 1.74, SD = 1.17), the race of the suspect (M = 1.57, SD = 1.00), and the race of the witness (M = 1.39, SD = .84). I performed a factor analysis on the influence factors for further investigation and found three factors: Experimenter Behaviour, Witness, and Race Information (see Table 3).

I performed a MANOVA to examine the impact of Suspect Race (No Information v. White v. South Asian) on the scaled self-report influencing Factors (Experimenter behaviour v. Witness behaviour v. Race/Appearance). I did not include witness race because it was a within-subject factor. The impact of suspect race was not significant for Experimenter factors, $F(2,156) = .451$, $p = .638$, Witness behaviour ratings, $F(2,156) = .464$, $p = .630$, or Race information, $F(2,156) = 1.648$, $p = .097$ (see Figure 5). To investigate whether the ratings differed significantly across factors, I conducted a series of paired t-tests and found that witness behaviour (M = 4.12, SD = .65) was rated as significantly more influential than Experimenter Factors (M = 3.51, SD = 1.13), $t(158) = 6.472$, $p < .001$, and Race Information (M = 1.57, SD = .84), $t(158) = 27.780$, $p < .001$. The mean for Experimenter Factors was significantly higher than for Race Information, $t(158) = 17.107$, $p < .001$.

When asked whether they agreed with the statement that other-race identifications are harder to make than same-race identification, the majority of participants completely disagreed (32.1%) or were neutral (30.2%). In order to test whether participants who thought other-race identifications are harder had more difficulty with them, I computed a correlation between accuracy score for other-race identifications and agreement with this statement. The correlation was not significant, $r = -.049$, $n = 105$, $p = .621$. This means that participants who agreed that other-race identifications are harder to make then same-

race identifications and participants who disagreed with this statement were about equal in their abilities to discriminate accurate from inaccurate other-race identifications.

Discussion

Summary of Results

Researchers have found that witnesses are more accurate at identifying people of their own races than of other races (Meissner & Brigham, 2001). I examined peoples' abilities to discriminate accurate from inaccurate same- versus other-race identifications. I hypothesized that I would find that same-race identifications would be believed more than other-race identifications, but that participants would be no more accurate in evaluating the accuracy of same-race than other-race identifications. I found that participants perceived the White witnesses as more accurate as well as more credible than the South Asian witnesses, regardless of suspect race information.

With respect to actual accuracy, I hypothesized and found that, overall, participants were unable to discriminate accurate from inaccurate identifications beyond chance levels. Also, as I hypothesized, participants were unable to discriminate accurate from in accurate same-race identifications. Contrary to my hypothesis, participants were able to discriminate accurate from inaccurate other-race identifications at a rate significantly above chance (57%), though accuracy was still weak.

I hypothesized that participants would be no better at discriminating the accuracy of same- and other-race identifications. My hypothesis was not supported, and the results were complex. Although participants performed similarly across conditions when viewing White Witness videos, the results fluctuated when viewing South Asian witness videos. When the suspect was thought to be White, participants were significantly better able to distinguish between accurate and inaccurate South Asian witnesses than when they were given no suspect race information or when the suspect was South Asian.

The confidence levels of participants did not vary significantly across conditions. Participant gender and ethnicity also did not affect their perceived or actual accuracy judgements. The final set of results, pertaining to the self-report questions at the end of the study, indicated that participants reported that witness behaviour factors were the most influential, significantly more so than experimenter factors and race information. Participants rated race information as significantly less of an influence than either of the other two factors.

Why were White witnesses perceived to be more accurate than South Asian witnesses? I re-examined the videos used in the study and compared the time it took for each witness to make an identification. I found that the White witnesses (M = 1.50 sec., SD = 1.27) did not make identifications significantly faster than did South Asian witnesses (M =3.89 sec, SD = 4.73), $t(17) = -1.54$, $p = .14$. I also examined whether White witnesses gave higher confidence ratings following their identifications as compared to South Asian witnesses. Mean confidence for White witnesses (M = 7.90) did not differ significantly from mean confidence for South Asian witnesses (M = 7.61), $t(17) = .26$, $p = .80$. Perhaps participants perceived the White witnesses in the videos as more confident than the South Asian witnesses, even if the witnesses' own verbal expressions of confidence did not differ significantly between races. The higher perceived accuracy of White witnesses does not seem to be related to witness expressed confidence, but rather perceived credibility. Perhaps people have a negative bias towards South Asian witnesses, and these stereotypes could lead to participants believing they are less accurate and less credible than the White witnesses.

With respect to actual accuracy, I found that participants were better able to discriminate between accurate and inaccurate White witnesses than South Asian witnesses. Finding that participants were more accurate in distinguishing between accurate and inaccurate perceived other-race identifications than same-race identifications (but only with the South Asian suspect) goes against previous research and assumptions (Abshire & Bornstein, 2003; Meissner & Brigham, 2001). Perhaps White witnesses in the videos displayed more behavioural cues associated with accuracy. For example, the White witnesses may have sounded more confident in their identification when they were correct, even if verbal ratings of witness confidence were not significantly different between the races.

Although witness race affected perceived accuracy and actual accuracy of the identifications, participants reported that the races of the suspects and witnesses had the least amount of influence on their judgements compared to the experimenter factors and witness behaviour. Perhaps participants were trying to be socially desirable by saying that race did not influence their judgements when, in reality, the participants may find White witnesses to be more credible and persuasive than South Asian witnesses. Past research has found that global assessments of people can be related to their perceptions of a person's mannerisms, demeanor, and other attributes (Nisbett & Wilson, 1977). Often, people are unaware of the factors that affect their decisions. Nisbett and Wilson (1977) found that the factors people believe they are relying on do not seem to affect their judgements, and the factors that they believe do not affect their judgements in fact do. With these findings in mind, it is not surprising that the participants' perceptions of the

role of suspect and witness race on their judgements did not match the effect that these factors actually had.

Comparison to Previous Research

My study makes unique contributions to the literature in several respects. Unlike Martire and Kemp (2009), who found that participants performed better than chance level when evaluating the accuracy of eyewitnesses, I found that participants performed below chance level when evaluating South Asian witnesses but at chance level for same-race identifications, for each of the suspect race conditions, and for overall accuracy. Participants performed better than chance when evaluating White witnesses and other-race identifications I also found that participants were more accurate in the other-race condition than in the same-race condition when the suspect was portrayed as South Asian. Perhaps the presence of a minority suspect and a White witness made participants pay closer attention to the video and thereby improve their accuracy. It is a puzzling finding, and more research would be needed to see if the finding replicates and, if so, to help explain it.

Abshire and Bornstein (2003) found that participants failed to take cross-race information into account when making accuracy judgements even when they are aware of the cross-race effect. A limitation of that study, however, was that the researchers only investigated the cross-race effect with a Black defendant and could only assess perceived accuracy scores. My design enabled me to assess both White and South Asian witnesses and defendants, as well as both perceived and actual accuracy of same- and other-race identifications, with different results. I did not find significant results for participant gender and ethnicity. Abshire and Bornstein found that participants were more likely to

believe the witness in same-race identifications (Black witness, Black suspect) than in other-race identifications (White witness, Black suspect). Interestingly, I found that participants were more likely to believe White witnesses over South Asian witnesses, but were more accurate in other-race identification than same-race identifications (when the suspect was South Asian). By examining both White and South Asian witnesses and defendants, I was able to build on the research and generalizability claims of Abshire and Bornstein (2003). A second contribution I made through my research is the ability to analyze both perceived and actual accuracy. It is important to look at both perceived and actual accuracy because they are two different concepts, and both are important in actual jury settings. In jury trials, juror belief of a witness is what matters. The problem is, that witness could be correct or they could be incorrect in their identification and we would not know. Analysis of actual accuracy tells us more about people's cognitive abilities to evaluate eyewitnesses. By investigating both perceived and actual accuracy rates, I was able to get a better sense of the relationship between the two concepts in other race situations, and notice that perceived accuracy is not always indicative of actual accuracy.

Limitations

In my attempt to address some limitations of past research, I inevitably encountered limitations in my own research. One limitation of my research is the sample of videos (20) watched by participants. A larger, more balanced video selection for both White and South Asian witnesses may have produced results more consistent with my hypotheses as well as previous research. Although the confidence ratings of the witnesses did not vary across videos, it is possible that there were subtle differences in witness behaviour across races, and this could have impacted participants' abilities to judge

accuracy. There may have been subtle cues to accuracy or inaccuracy present in the videos for one race and not the other race. Future research should code the behaviours in the videos and attempt a more balanced video sample. These results could be unique to this set of videos and may not generalize to other cross-race identification situations. More research needs to be conducted to see if my results can be replicated with a completely different set of videos. A second limitation is that I did not have true other-race identifications for South Asian suspects, as I manipulated the suspect race with a photo after the video was filmed. In the original study in which the videos were gathered, these would have been same- and cross-race identifications, because there were White and South Asian witneses and White suspects, but the design would be imbalanced because there was no South Asian suspect. By manipulating the race of the suspect, I was able to create the illusion of other-race identifications; however the artificiality of my manipulation may have impacted my results. The behaviours exhibited by the participants would have been the same across suspect race conditions (no information, White suspect, South Asian suspect), but the presence of South Asian suspects could cause witnesses to behave differently. A final limitation is the population from which my sample of participants was recruited. A number of researchers have suggested that undergraduate students may not be a representative sample of the community or population in general (Bornstein, 1999); therefore, a sample that included both undergraduate students and a variety of community members may make my results more reliable and generalizable. As the minimum age for a person to be selected jury duty is 18 years, my participants were on the very low end of the age range. Chances are the mean age for a typical jury will be greater than the mean age for my participants, and this

could greatly change the results. More life experience, typically associated with older ages, may involve more interactions with people of other races, changing their perceptions and racial attitudes (Meissner & Brigham, 2001). If this is the case, then younger jurors, around the mean age of my jurors, may have different racial attitudes owing to less interaction with races other than their own. Future research should examine whether results of my thesis generalize to other stimuli and samples.

Summary and Conclusions

Over the past four decades, numerous researchers have attempted to understand eyewitnesses, various factors affecting eyewitness identifications, and juror evaluations of these eyewitness identifications (e.g., Steblay et al., 2001; Steblay, et al., 2003; Yarmey, Yarmey, & Yarmey, 1996). This research is important because of the weight placed on eyewitnesses in the criminal justice system. Unfortunately, eyewitnesses make mistakes (Steblay et al., 2001; Steblay, et al., 2003; Yarmey et al., 1996). The main goal of my thesis was to gain further insight into the factors that influence peoples' abilities to discriminate accurate from inaccurate eyewitness identifications, namely the race of the eyewitness and race of the suspect. I found that White witnesses were viewed as more accurate and credible than South Asian witnesses. In addition, participants were better able to distinguish between accurate and inaccurate South Asian witnesses when the participants viewed what they perceived to be other-race identifications than when they viewed same-race identifications. In the future, it would be beneficial to code the videos used in my research for behavioural cues present in the witness videos that participants judged correctly. Further research should also examine other factors that may affect jurors' abilities to discriminate accurate from inaccurate eyewitnesses.

References

Abshire, J. & Bornstein, B. H. (2003). Juror sensitivity to the cross-race effect. *Law and Human Behaviour, 27,* 471-480. doi:10.1023/A:1025481905861.

Beaudry, J. L. (2008). *Video-recorded lineup procedures and detecting identification accuracy* (Doctoral dissertation). Retrieved from ProQuest. (NR44310). Queen's University, Kingston, ON, Canada.

Bornstein, B. H. (1999). The ecological validity of jury simulations: Is the jury still out? *Law and Human Behavior, 23,* 75-91. doi:10.1023/A:1022326807441.

Cutler, B. L., & Penrod, S. D., & Dexter, H. R. (1990). Juror sensitivity to eyewitness identification evidence. *Law and Human Behaviour, 14,* 185-191. doi:10.1007/BF01062972.

deLone, M. (2011). Looking back, looking ahead. *The Innocence Project In Print, 7,* 3.

Desmarais, S. L. & Read, J. D. (2011). After 30 years, what do we know about what jurors know? A meta-analytic review of lay knowledge regarding eyewitness factors. *Law and Human Behavior, 35,* 200-210. doi:10.1007/s10979-010-9232-6.

Huff, C. R. (2004). Wrongful convictions: The American experience. *Canadian Journal of Criminology and Criminal Justice, 46,* 107-120. doi:10.3138/cjccj.46.2.107.

Lindsay, R. C. L., Wells, G. L., & O'Connor, F. J. (1989). Mock-juror belief of accurate and inaccurate eyewitnesses: A replication and extension. *Law and Human Behavior, 13,* 333-339.

Martire, K. A. & Kemp, R. I. (2009). The impact of eyewitness expert evidence and judicial instruction on juror ability to evaluate eyewitness testimony. *Law and Human Behaviour, 33,* 225-236. doi:10.1007/s10979-008-9134-z.

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,* 3-35. doi:10.1037//1076-8971.7.1.3.

Nisbett, R. E. & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35,* 250-256.

Read, J. D. & Desmarais, S. L. (2009). Lay knowledge of eyewitness issues: A Canadian evaluation. *Applied Cognitive Psychology, 23,* 301-326. doi:10.1002/acp.1459.

Smith, A. M., Leach, A.-M. & Cutler, B. L. (2011, March). Showup identifications: The effects of presence of stolen property and suspect's denial on identification performance. Paper presented at the American-Psychology-Law Society Annual Conference, Miami, FL.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2001, October). Eyewitness accuracy rates in sequential and simultaneous lineup presentations: A Meta-analytic comparison. *Law and Human Behaviour, 25,* 459-473. doi:10.1023/A:1012888715007.

Steblay, N., Dysart, J., Fulero, S., & Lindsay, R. C. L. (2003). Eyewitness accuracy rates in police showup and lineup presentation: A meta-analytic comparison. *Law and Human Behaviour, 27,* 523-540. doi:10.1023/A:1025438223608.

Wells, G. L., Lindsay, R. C. L., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64,* 440-448. doi:10.1037/0021-9010.64.4.440.

Yarmey, D. A., Yarmey, M. J., & Yarmey, L. A. (1996). Accuracy of eyewitness identification in showups and lineups. *Law and Human Behaviour, 20,* 459-477. doi:10.1007/BF01498981.

Table 1

Descriptive Statistics for Participant Ability to Judge Witness Accuracy

Variable	N	*M*	*SD*
Accuracy by Condition			
No Suspect Race Information	54	.49	.06
White Suspect	55	.52	.09
South Asian Suspect	50	.50	.08
Accuracy by Witness Race			
White Witnesses	159	.57*	.12
South Asian Witnesses	159	.43*	.11
Accuracy by Identification Type			
Same-Race Identification	105	.48	.15
Other-Race Identification	105	.53*	.13
Overall Accuracy	159	.50	.08

*$p < .05$. Chance is 50%

Table 2

Descriptive Statistics for the Influential Factors, Organized by Factor

Influential Factor	N	*M*	*SD*
Experimenter Factors	159	3.51	1.13
How the experimenter behaved	159	3.50	1.28
What the experimenter did/did not say	159	3.52	1.26
Witness Behaviour	159	4.12	.65
How the witness behaved	159	4.45	.85
The confidence of the witness	159	4.58	.84
How long it took the witness to make an ID	159	4.09	.95
How carefully the witness looked at the suspect	159	3.72	1.31
How nervous the witness appeared	159	3.69	1.10
What the witness did/did not say	159	4.21	.87
Race Information	159	1.57	.84
The race of the suspect	159	1.57	1.00
What the witness looked like	159	1.74	1.17
The race of the witness	159	1.39	.84

Table 3

Factor Loadings for Exploratory Factor Analysis with Varimax Rotation of Influential Factors Scales

Scale	Experimenter Factors	Witness Behaviour	Race Information
How the experimenter behaved	**.869**	.073	.034
What the experimenter did/did not say	**.884**	.083	-.032
How the witness behaved	.005	**.581**	-.516
The confidence of the witness	.072	**.688**	-.335
How long it took the witness to make an ID	.124	**.736**	-.050
How carefully the witness looked at the suspect	-.031	**.686**	.098
How nervous the witness appeared	-.017	**.677**	.120
What the witness did/did not say	.331	**.597**	-.163
The race of the suspect	.101	-.106	**.799**
What the witness looked like	-.126	.049	**.791**
The race of the witness	-.003	-.017	**.863**

Note. Factor loadings > .50 are bolded.

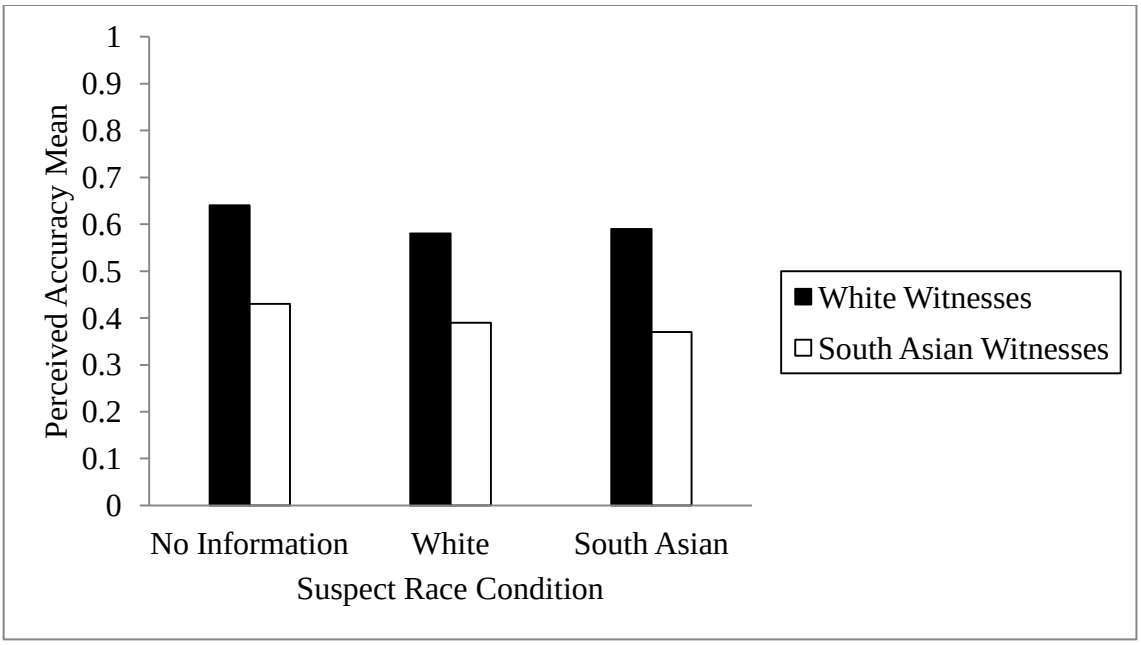

*Figure 1.*Perceived accuracy means for White and South Asian witnesses as a function of suspect race.

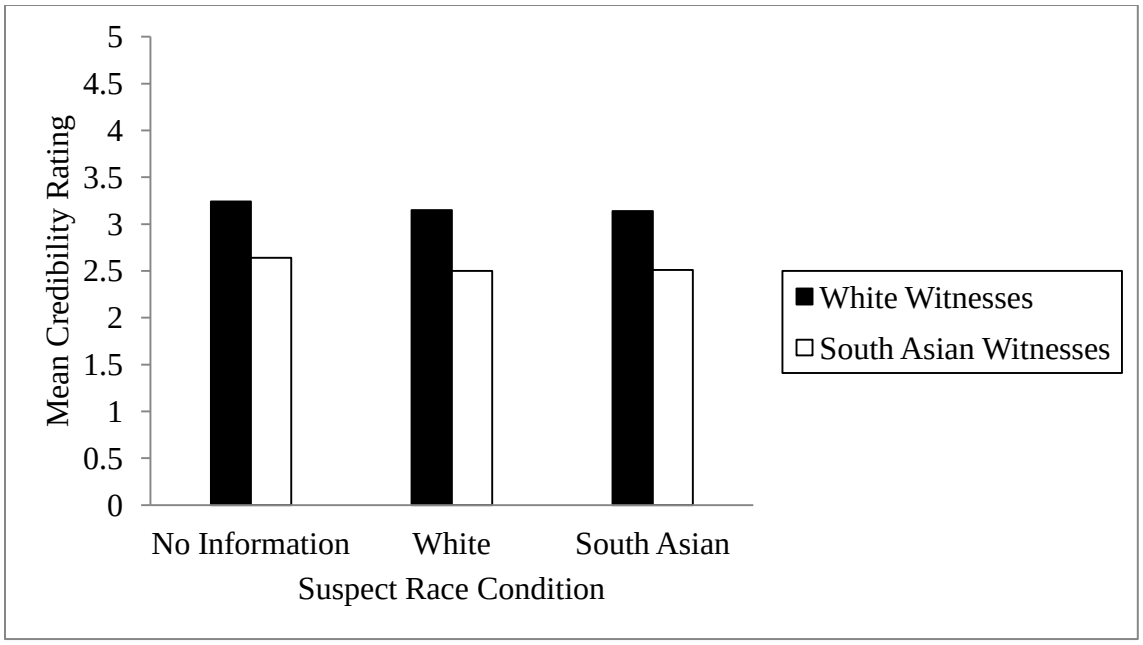

Figure 2. Credibility rating means for White and South Asian witnesses as a function of suspect race.

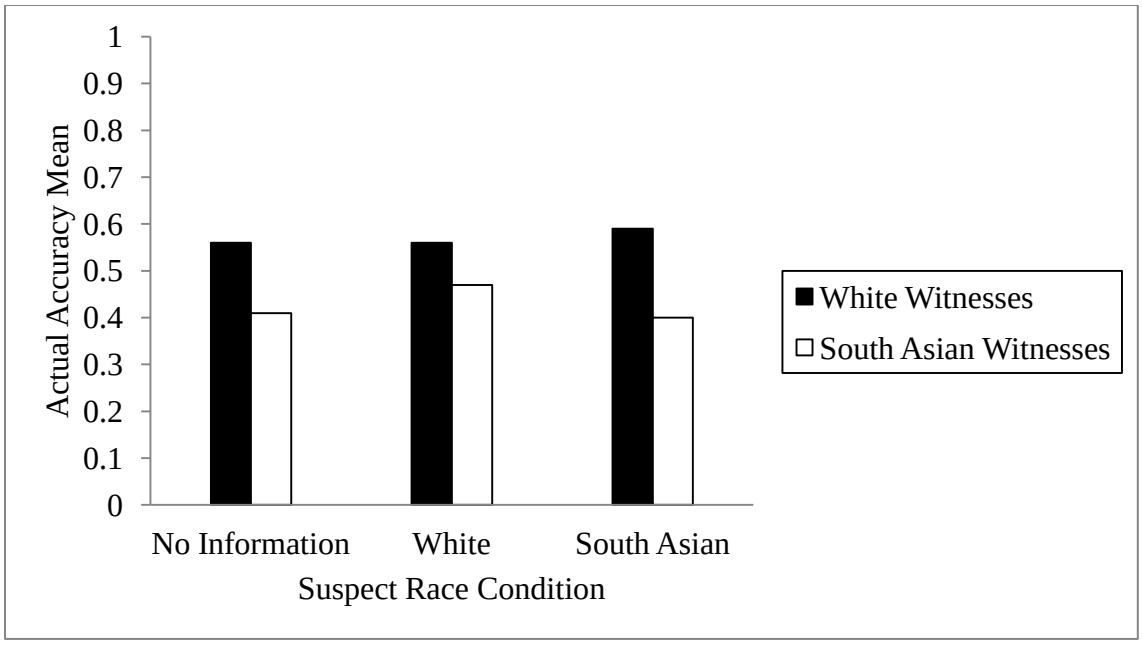

*Figure 3.*Actual accuracy for White and South Asian witnesses as a function of suspect race.

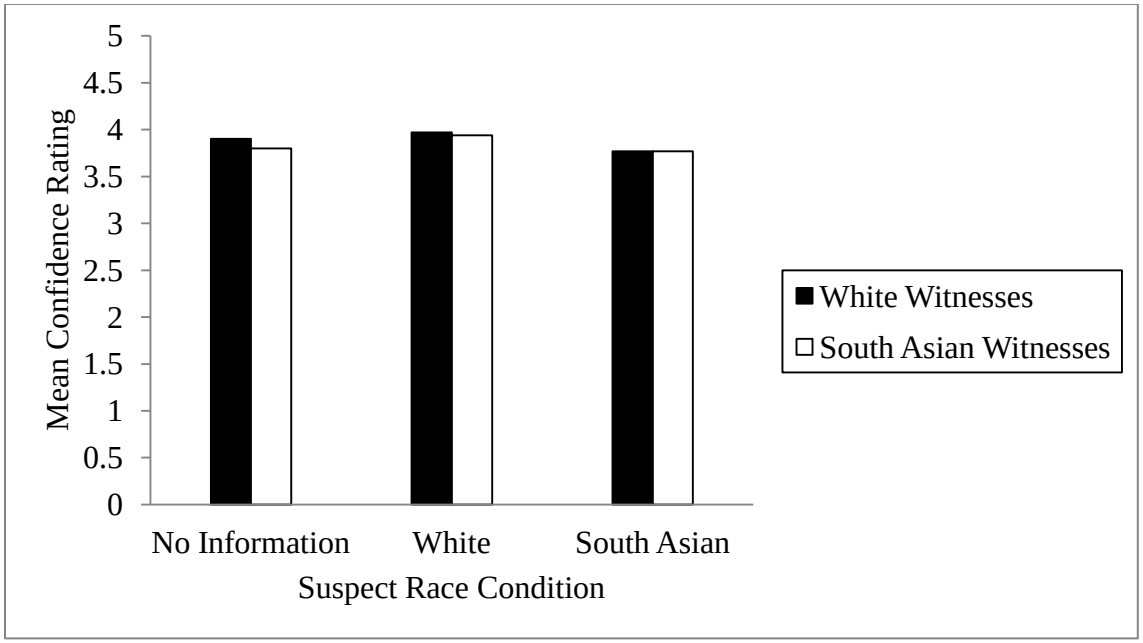

Figure 4. Confidence rating means for White and South Asian witnesses as a function of suspect race.

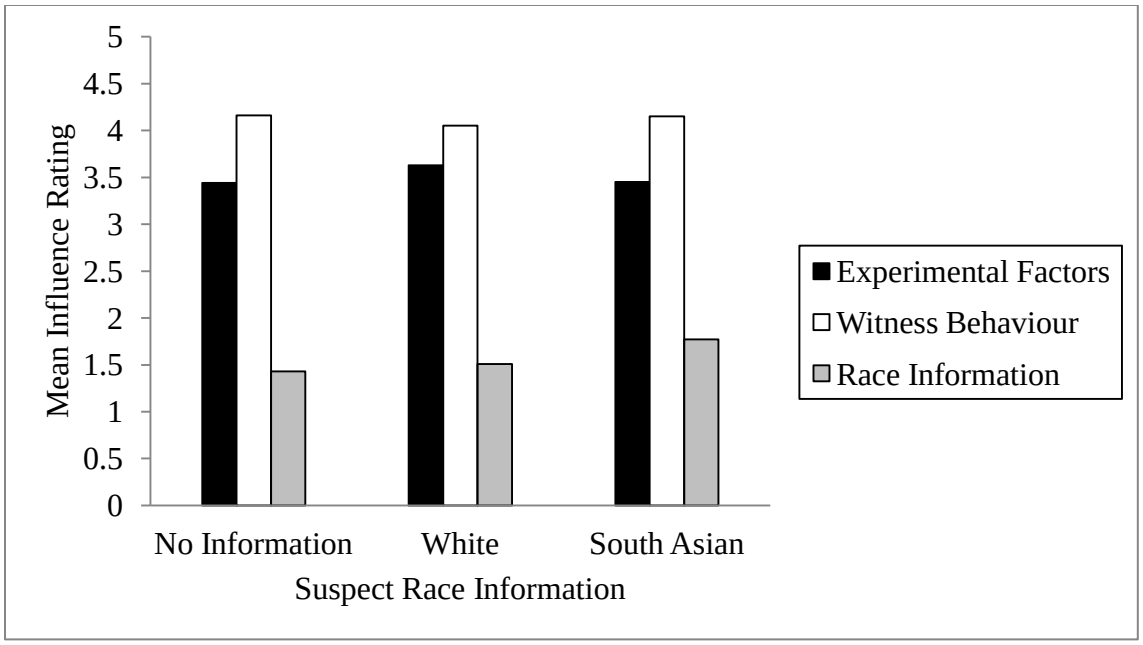

Figure 5. Influential Factors means for the three factors as a function of suspect race.

Appendix A – Ethics Approval

Date: November 3rd, 2011
To: Katherine Reiss (Student PI), Brian Cutler (Supervisor)
From: Shawn Bullock, REB Vice - Chair
REB File #: 11-043
Project Title: Investigating the Ability to Judge the Accuracy of Eyewitnesses
DECISION: APPROVED
START DATE: November 3rd, 2011 EXPIRY: November 3rd, 2012

The University Of Ontario Institute Of Technology Research Ethics Board has reviewed and approved the above research proposal. The application in support of the above research project has been reviewed by the Research Ethics Board to ensure compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS2) and the UOIT Research Ethics Policy and Procedures. Please note that the Research Ethics Board (REB) requires that you adhere to the protocol as last reviewed and approved by the REB.

Always quote your REB file number on all future correspondence.
Please familiarize yourself with the following forms as they may become of use to you.

- **Change Request Form:** any changes or modifications (i.e. adding a Co-PI or a change in methodology) must be approved by the REB through the completion of a change request form before implemented.
- **Adverse or unexpected Events Form:** events must be reported to the REB within 72 hours after
the event occurred with an indication of how these events affect (in the view of the Principal Investigator) the safety of the participants and the continuation of the protocol. (I.e. unanticipated or un-mitigated physical, social or psychological harm to a participant).
- **Research Project Completion Form:** must be completed when the research study has completed.
- **Renewal Request Form:** any project that exceeds the original approval period must receive approval by the REB through the completion of a Renewal Request Form before the expiry date has passed.

All Forms can be found at http://research.uoit.ca/EN/main/231307/Research_Forms.html.

REB Vice - Chair
Dr. Shawn Bullock
shawn.bullock@uoit.ca

Ethics and Compliance Officer
Sascha Tuuha, (905) 721-8668 ext. 3693
compliance@uoit.ca

Appendix B – Questions Between Videos

I think the identification was:

1 Correct
2 Incorrect

On a scale from 1 to 5, how confident are you in your decision?

1 Not at all confident
2 –
3 –
4 –
5 Very confident

On a scale from 1 to 5, how credible do you find the witness?

1 Not at all credible
2 –
3 –
4 –
5 Very credible

Appendix C – Influencing Factors Questionnaire

In general, when you made your decision regarding whether the witness was correct or incorrect, what factors influenced those decisions? Please rate the extent to which each of the following factors influenced your decision on a scale from 1 (no influence) to 5 (very influential).

a. How the experimenter behaved:
b. What the experimenter did or did not say:
c. How the witness behaved:
d. The race of the suspect:
e. The confidence of the witness:
f. What the witness looked like:
g. How long it took the witness to make an identification:
h. How carefully the witness looked at the suspect:
i. The race of the witness:
j. How nervous the witness appeared:
k. What the witness did or did not say:

Do you agree or disagree with the following statement:

> "If the race of the suspect is different from the race of the witness, it is more difficult for the witness to make a correct identification."

1. Completely Disagree
2. –
3. Neutral
4. –
5. Completely Agree

Appendix D – Debriefing Form

You have just participated in a research project about eyewitness memory and juror judgement accuracy entitled Investigating the Ability to Judge the Accuracy of Eyewitnesses. You have viewed videoclips of eyewitnesses making identifications from show-ups, or field identification tests. Some participants watched witnesses identify people of their own race. Other participants watched witnesses identify people of another race. Other participants did not know the race of the suspects being identified. In each case some of the identifications were correct and some were incorrect. We are interested in learning whether the race of the witness and the race of the suspect, alone or in combination, influences how people such as yourself evaluate the accuracy of eyewitness identifications.

It is very important that you keep the true nature of this study confidential. For the study to be valid, participants must not know at the start of the study that we are interested in how witness and suspect race influence evaluations of eyewitness identification. Please assist us in maintaining the validity of the study by not discussing its purpose with anyone.

If you have any questions, please feel free to ask the researcher present, contact the principal investigator (Katherine.Riess@uoit.ca), her supervisor (Brian.Cutler@uoit.ca), or the Ethics and Compliance Officer (compliance@uoit.ca). Thank you very much for participating in the study.

Appendix E – Informed Consent Form

You have been invited to voluntarily participate in a research project about eyewitness memory entitled "Investigating the Ability to Judge the Accuracy of Eyewitnesses." We are interested in learning about how people evaluate eyewitnesses. You will be watching a series of video clips and answering questions regarding what you saw in the videos, as well as some generic demographic questions (e.g. your age, ethnicity, and faculty). The entire process should last about one hour.

There are no known physical, psychological, economic, or social risks associated with this study. Your participation in this study is completely voluntary and you may withdraw from this study at any time without any consequences or penalties. By participating, you are not giving up any legal rights. You are not obliged to answer any questions that you find objectionable or which make you uncomfortable. You will be given one full credit for your participation in this study. Full credit will be awarded whether you complete the study or not.

All information will be coded and stored in a secure area. Only the primary researcher, her assistants, her supervisor Dr. Cutler, and other researchers interested in psychology will have access to the data (e.g., for meta-analyses). Individual performance will remain confidential and will not be released to professors, employers or in publications. Only group results will be reported (e.g., conferences presentations, journal articles).

This study has been reviewed and cleared by the Research Ethics Board at UOIT (REB # 11-043). The principal investigator is Katherine Riess, a graduate student in the Faculty of Social Science and Humanities, UOIT. She is being supervised by Dr. Brian Cutler. In the event that you have any questions, concerns, or complaints, you may contact any of the following individuals: Dr. Cutler (brian.cutler@uoit.ca; 905-721-8668 ext. 3807) or the Ethics and Compliance Officer (compliance@uoit.ca; 905-721-8668, ext. 3693).

I have read and understood the statements above. I have had my questions answered to my satisfaction and I understand that I may ask additional questions at any time. My signature, below, indicates my free and informed consent to participate in this research.

Name (please print) _____
Signature _____
Date _____

Appendix F – Demographic Questions

I. Please state your age.

II. Please indicate your gender:
   1 Male
   2 Female

III. What is your ethnic origin?
   1 Aboriginal (Inuit, Metis, North American Indian)
   2 Arab/West Asian (e.g. Armenian, Egyptian, Iranian, Lebanese, Moroccan)
   3 Black (e.g. African, Haitian, Jamacan, Somali)
   4 Caucasian (White)
   5 Chinese
   6 Filipino
   7 Hispanic
   8 Japanese
   9 Korean
   10 South Asian
   11 South East Asian
   12 Other (Specify)

IV. What faculty are you in?
   1 Faculty of Business and Information Technology
   2 Faculty of Education
   3 Faculty of Engineering and Applied Science
   4 Faculty of Energy Systems and Nuclear Science
   5 Faculty of Health Sciences
   6 Faculty of Science
   7 Faculty of Social Science and Humanities

V. What year of your program are you in?
   1 First Year
   2 Second Year
   3 Third Year
   4 Fourth Year
   5 Other (Specify)