That's the Guy!

The Impact of Eyewitnesses' Prejudice on the Composite Production of Perpetrators' Faces

Jascha Bentlage

Faculty of Behavioral Sciences, University of Twente

Bachelor thesis

1st supervisor: Dr. Elze Ufkes

2nd supervisor: Dr. Ir. Peter W. de Vries

Specialization: Psychology of Conflict, Risk & Safety

Twente University

Date: 24 June 2015

Abstract

Eyewitness misidentifications are the main reason for wrongful convictions. Most victims of judicial errors belong to minority groups. This study consists of two parts. In Study 1 participants completed an implicit association test to measure their level of prejudice. Thereafter, they constructed composites of either an in-group or an out-group perpetrator using the reverse-correlation image classification technique (N = 21). In a follow-up study those produced pictures were rated by independent individuals (N = 116) to measure their level of psychological valence. It was expected that pictures of out-group faces produced by participants with high level of prejudice would be rated as more negative looking. Results showed that highly prejudiced participants produced more negative looking stimuli for both the in-group and for out-group faces. These findings counteract the hypothesis that the effect of prejudice would only be observable for out-group faces. Recommendations for further research are suggested.

Introduction

It was a normal forenoon in the year of 1982 when three men trespassed into the house of a young couple in Dallas, Texas. They robbed the man and forced his pregnant wife to come with them. The 18 year old woman had to undress herself on a meadow and was raped by the three men. One month later the police received an anonymous hint: James Giles was one of the perpetrators. In the following line-up the woman identified him as well. Giles denied it vehemently, but was still sentenced because of brutal rape to jail for life. It took 25 years until Giles was acquitted on basis of a DNA test (Blech, 2007).

Unfortunately, the case of James Giles cannot be seen as one horrible exception. Until now, 325 men were released out of prison because of DNA tests which proved their innocence. 72 % of them were convicted because of eyewitness misidentification, which is by far the main reason for judicial errors like this, and 70 % of them belong to minority groups, the majority of them have an Afro-American background (Innocence Project, n.d.). Considering these facts, two questions come to mind: why are people obviously relatively bad at recognizing perpetrator's faces and to what extent does the racial background of the suspect play a role? This study investigates in how far the level of prejudice of an eyewitness influences the perceived psychological valence of the faces he or she produces.

Reliability of composite systems

Research shows that people's ability to produce intended faces from memory is generally bad. This is true for traditional pen and paper versions for composite systems (Davies, Ellis, & Shepherd, 1978) as well as for modern versions using computers (Davies & Valentine, 2006). Studies using recognition produce generally better results than studies using recall systems (Davies & Donald, 1982). But, considering the importance eyewitness's testimonies may have,

also systems making use of recognition still produce insufficient results. A possible explanation for the bad performance is the way in which human beings process, store, and retrieve faces. This happens more at a holistic rather than a feature-based level (Tanaka & Farah, 2003). This means, that the "human visual system does not process a face as a collection of separable facial features but as an integrated whole" (Taubert, Apthorp, Aagten-Murphy, & Alais, 2011). The method of face-composite systems counteracts this process because eyewitnesses have to select individual features. Thereby comparisons between different parts of the face have to be made which may also interfere with the representation of the target face (Wells & Hazel, 2007).

Recent research has tried to develop a better functioning system for the identification of perpetrators. Bruce and colleagues (2002) found that the method of morphing may be promising, although the results are still poor. Several faces of possible perpetrators are simultaneously presented to the eyewitness. He has to choose which picture fits the representation of the face he holds in memory at best. Afterwards, the selections are put together into one composite by aggregating at the pixel level. The results of Bruce et al (2002) were supported by another study, which found that morphs were rated as more similar to the target face than those of individual composites (Wells & Gary, 2007).

For this reason, the present study makes use of the reverse-correlation image classification technique which is based on recognition and follows the approach of holistic face representation (Mangini & Biedermann, 2004). Faces of perpetrators are created by adding different levels of noise to a base face. Participants have to choose between different faces which are simultaneously presented. Afterwards, the selections are aggregated and morphed into a single composite.

Influence of prejudice

Social factors may play an important role at recognizing perpetrators in a judicial setting. It is a well-documented fact that new faces are automatically categorized into social groups (Alport, 1954). Faces of out-group members are processed faster than those of in-group members but the accuracy suffers (Levin, 1989; Valentine & Endo, 1992). A study by Eyseneck and Keane (2013) which examined more than 200 real court cases showed that eyewitnesses correctly identified 65 % of their in-group while accuracy for out-group members was only 45 %. This cross-race effect was replicated in a large number of studies (Anthony, Cooper, & Mullen, 1992; Chance & Goldstein, 1981; Cross, & Daly, 1971; Malpass & Kravitz, 1969).

One possible explanation for the cross-race effect is the human tendency to attribute different degrees of humanity to different social groups. Harris and Fiske (2006) showed in their study using functional magnetic resonance imaging that the medial prefrontal cortex is not activated when participants thought of out-group members. This area is usually active while thinking of persons, but not of objects. Thus, out-group members may be "inhumanized" and may be seen more as objects rather than human beings. Those findings were supported by another study which showed that people tend to protect the human integrity of their in-group by assigning signs of inhumanity only to out-group members, but not to in-group members (Capozza et al, 2009). Bernstein et al. (2007) investigated whether this effect is only observable for already existing groups, thus predominantly ethnicities, or also for other social categories. Therefore they experimentally created minimal groups. The results showed that merely social categorization is sufficient to influence the accuracy of recognition. This finding is also supported by another study (Sangrigoli et al, 2005), which showed that Asian children who were adopted and raised in a Western environment had a better recognition for Caucasian than for Asian faces. Thus, the cross-race effect is not only a result of genetic heritage, but changeable due to environmental

circumstances. Social categorization obviously plays a role in the recognition of faces, especially in situations in which individuals have to identify out-group members.

But, what does all this have to do with eyewitness identification in a judicial setting? The pre-existing beliefs an investigator or member of a jury holds, may influence his or her decisions regarding the guilt of a suspect (Holyoak & Simon, 1999; Simon, Pham, Le, & Holyoak, 2001; Simon, Snow, & Read, 2004). Thus, one's ethnic background can possibly be harmful for suspects. In the Netherlands, Moroccans are a highly stigmatized group and have to struggle with prejudice towards them (Coenders, Lubbers, Scheepers, & Verkuyten, 2008; Verkuyten, 2008; Verkuyten & Zaremba, 2005). Prejudice may be defined as a "feeling, favorable or unfavorable, toward a person or thing, prior to, or not based on, actual experience." (Allport, 1954). The majority of the Dutch people hold negative stereotypes about the Moroccans and associate them with the trait 'criminal' (Gordijn, Koomen, & Stapel, 2001). Previous research has shown that expectations and beliefs a composite-creator holds about a target influence the composites he or she creates (Davies & Oldman, 1999; Shephered, Ellis, McMurran, & Davies, 1978). In a study by Dotsch et al (2008) Dutch participants had to construct such composites of faces of which they believed a typical Moroccan looks like. Thereafter, they conducted an implicit association test (IAT). This test is suited to measure the extent to which people differ in their level of implicit prejudice (Greenwald, McGhee, & Schwartz, 1998). The pictures were produced using a force choice paradigm, resulting in composites of different faces. Composites composed by highly prejudiced participants were rated as more criminal-looking than faces produced by individuals low in prejudice.

Individuals vary in their level of prejudice and more prejudiced persons produced more aggressive looking pictures. Similar findings have been established by several studies (Allport & Kramer, 1946; Elliot & Wittenberg, 1955; Himmelfarb, 1966; Quanty, Keats, & Harkins, 1975;

Kosic & Phalet, 2006; Castano, Yzerbyt, Bourguignon, & Seron, 2002). Also, the finding that individuals vary in their level of prejudice is supported by earlier studies (Gordijn, Koomen, & Stapel, 2001; Wittenbrink, Judd, & Park, 1997). Moreover, Dotsch and colleagues (2011) did not test for in-group perpetrators, but only made use of Moroccan out-group perpetrators. We found no study which directly compared the effect prejudice has on the identification between in-group and out-group perpetrators. Therefore, the present study aims to assess whether the effect of prejudice is only observable for out-group perpetrators, or also for Caucasian in-group perpetrators.

The current study asked participants to produce composites using the reverse-correlation image classification technique. This approach was successful in creating morphs which resembled intended character traits (Todorov & Dotsch, 2012). But, in contrast to Dotsch et al (2011) participants in the current study were provided with an image of a perpetrator. Thus, they were not asked to think of a typical Moroccan face, thereby directly activating their stereotype, but an actual picture of either a Dutch or a Moroccan perpetrator was given. Hence, the stereotype of the perpetrator is activated in a more indirect manner compared to Dotsch et al (2011). In this way, we want to assess if the type of stereotype activation plays a role. By using two conditions, we were able to examine whether the effect of prejudice was only observable for Moroccan out-group, or also for Caucasian in-group perpetrators. Later, the created morphs were rated in a follow-up study in terms of character traits like aggressiveness. In this way we wanted to assess whether the level of prejudice of a composite-constructor had an influence on the perceived psychological valence of the produced pictures. Psychological valence, in this context, is defined as the "intrinsic attractiveness or averseness" of the face (Frijda, 1998). Thus, the research question is: 'To what extent does the level of prejudice of the composite-constructor

influence the perceived valence of the produced picture?'. Considering the above mentioned findings, the hypothesis we tested was:

Composite-creators with high level of prejudice produce faces which will be labeled as more negative looking than those produced by less prejudiced individuals. This effect is only true for faces of out-group members.

To test this hypothesis, in the first experiment the level of prejudice of each participant was measured. Afterwards, she or he created a composite of either an in-group (Caucasian) or an out-group (Moroccan) face. Those composites were rated by independent individuals in a second study in terms of character traits like aggressiveness or friendliness to measure their perceived psychological valence.

Study 1 - Generating composites using reverse correlation

Method

Participants and design

Participants. In total, 22 students (9 male and 13 female) of the University of Twente participated in this study. Only the data of 21 of the respondents could be used for further analysis, because data storage for the other participant did not work. Of the 21 participants left, 9 were male and 12 were female. Five of them were Dutch and the other 16 were German. Age varied between 18 and 24 years with a mean score of 19,5 years. 18 of the students are studying

Psychology and the other 3 are studying Communication sciences at the University of Twente. In return, they received one course credit via SONA Systems or some confectionery.

Design. After completing the IAT, participants were randomly assigned to one of the following two conditions: an in-group target condition - with the perpetrator being Caucasian – or an out-group target condition – with the perpetrator being Moroccan. Thus, the experiment employed a 1-factorial between participants design, with the 2 conditions explained above. The dependent variable was the accuracy of the reconstructed face of the actual perpetrator.

Procedure

At their arrival, participants were directed to a small and quiet room in the library with a laptop in it. Before continuing with the experiment, researchers asked the participant for his or her informed consent. Furthermore, participants were instructed that it is important for the usability of the study that they stay focused, and that they are allowed to take a 5-minute break every 20 minutes to ensure this. The experiment was split into two different parts. First they had to complete an Implicit Association Test (IAT), so that afterwards we would be able to measure their prejudices and stereotypes towards Moroccans. Thereupon the reverse-correlation image classification task started, in which participants had to choose between 800 pairs of faces to reconstruct the offender they saw before.

Implicit prejudice. In order to measure prejudice, participants completed an Implicit Association Test (IAT). This test measured indirectly to what extent participants categorized Moroccan or Dutch faces into the positive or the negative group. We used a shortened version of the IAT, which consisted of a practice, congruent and incongruent block. In the practice block, participants had to classify 10 positive images with one key and 10 negative images with another. After that, participants had to complete 40 trials each in the congruent and in the incongruent

block. The order of the last two blocks was randomly distributed across participants. In the congruent block, participants classified Moroccan faces as a negative stimulus and Dutch faces as a positive one. In the incongruent block, participants did the reverse. Within blocks, stimuli were presented in a random order. When a mistake occurred during classification, error feedback was presented to the participant for 1,000 ms. If response latencies were longer in the incongruent block than in the congruent block, this was assumed to indicate stronger negative than positive associations with Moroccan faces. This difference was interpreted as reflecting higher levels of implicit prejudice. Implicit prejudice was measured using an implicit association test. Thereby, latencies above 3,000ms were set to 3,000ms. We analyzed the data on log-transformed latencies. We constructed an IAT score by subtracting the average response latency in the congruent block (M = 965.82, SD = 265.58) from the average response latency in the incongruent block (M = 1135.81, SD = 370.68).

Reverse-correlation image classification task. Participants were given further instructions about the following task. The system automatically assigned the participants randomly to either the in-group target or the out-group target condition. Two faces were selected from the Radboud Face Database (Langner et al., 2010), one typical Caucasian male face to represent the perpetrator in the in-group target condition, and a typical Moroccan male face for the out-group target condition. At the beginning of the task, participants were exposed with both a three-quarter view from both sides and a full-face view of the perpetrator. They were asked to study the face carefully by using the opportunity to switch between the different viewing angles. Subsequently, they were repeatedly presented with two stimuli side by side. Each pair of stimuli consisted of a base face with two random-noise patterns superimposed over it. The noise patterns were randomly generated at every one of the 800 trials. Participants were instructed to decide to what extent the stimuli resembled the target perpetrator shown in the beginning. They had 4 options to

choose from: 1) Clearly A, 2) Probably A, 3) Probably B and 4) Clearly B. Completing the 800 trials took the participants approximately 60 minutes.

Materials

Stimuli. As already mentioned above, participants repeatedly had to choose which of the base face with a random-noise pattern superimposed over it resembled the target more. In order to construct a base face, 12 faces have been selected from the Radboud Face Database (Langner et al., 2010): six Dutch and six Moroccan faces. We then morphed the features of those six faces into one base face for each condition (by using the program PsychoMorph). The noise pattern was generated by randomly calculating one set of parameters for each stimulus. Within a single trial, stimulus A consisted of the base face with a random-noise pattern, while the base face of stimulus B consisted of the inverse pattern (Figure X). The noise patterns distorted the base face to such an extent, that the pair of faces appeared to be different in every trial.

Results

Reverse-correlation image classification task. We used the script from Dotsch and Todorov (2012) to construct the classification images of the responses from the 21 participants for the second part of the study. The analysis yielded 11 classification images for the Dutch condition, and 10 classification images for the Moroccan condition. Furthermore, one final classification image was constructed for each condition by combining the features of all classification images into one average classification image (Figure 1).

FIGURE 1: FINAL CLASSIFICATION IMAGE FOR IN-GROUP (LEFT) AND OUT-GROUP (RIGHT) CONDITION

Implicit prejudice. The experiment used a one factor design, with ethnic background of the face as between-subject variable. Participants in the in-group condition produced faces of a Caucasian (Dutch) perpetrator, members of the out-group condition produced Moroccan faces. We used the script from Dotsch and Todorov (2012) to construct the classification images for the second study. Any incorrect trials and the first practice block were omitted from further analysis. A one-sample t-test revealed, that averaged, participants had more negative associations with the Moroccan category than positive ones (M = 1135.81, SD = 370.68), and stronger positive associations with the Caucasian category than negative ones (M = 965.82, SD = 265.58), $t(20)$= 2.67, $p< 0.001$. Using this data we calculated the median for both the Moroccan and the Dutch condition. By using the median split, we could identify participants that scored below or above the median on the implicit association test. If the IAT score in the Moroccan condition was above the median, this indicated a relatively stronger negative than positive association with Moroccan faces and is thus an expression of implicit prejudice. The same applied to the Dutch condition. This left us with a 2 x 2 design consisting of four cells: (1) out-group condition with low level of

prejudice, (2) out-group condition with high level of prejudice, (3) in-group condition with low level of prejudice and (4) out-group condition with high level of prejudice (Figure 2).

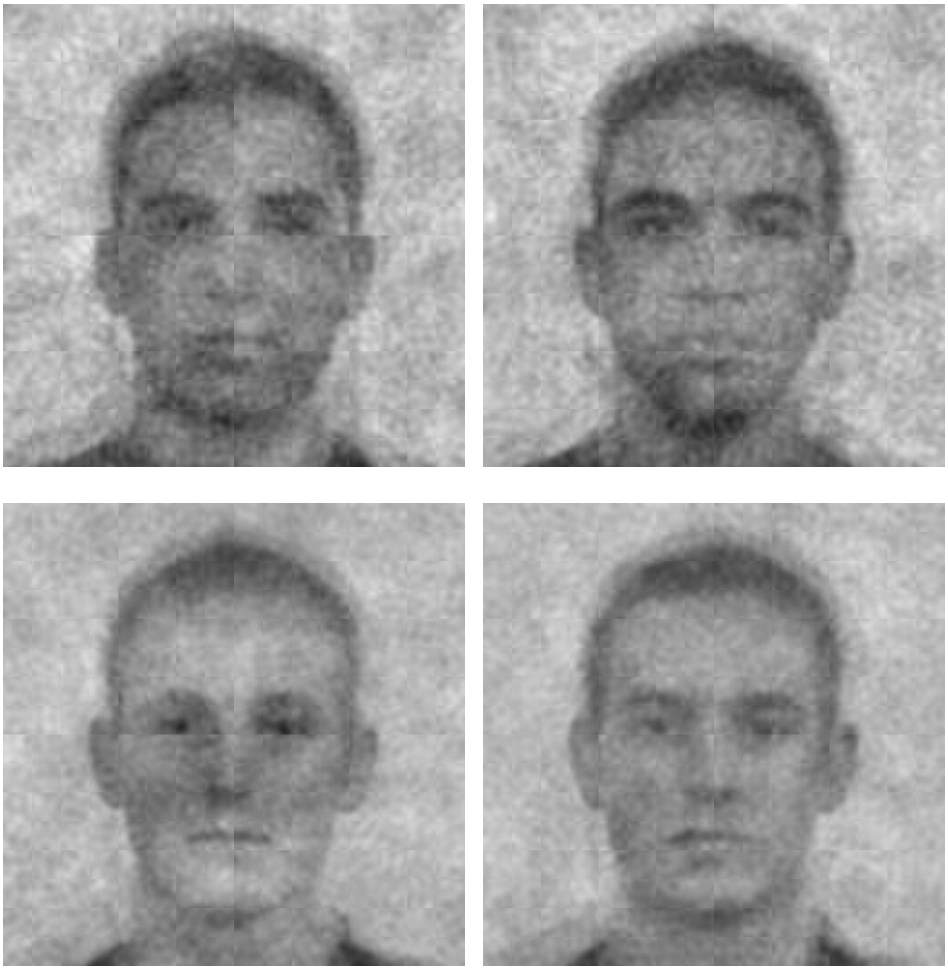

FIGURE 2: EXAMPLES FACES OF PERPETRATORS PER GROUP (TOP LEFT: OUT-GROUP WITH HIGH LEVEL OF PREJUDICE; TOP RIGHT: OUT-GROUP WITH LOW LEVEL OF PREJUDICE; BOTTOM LEFT: IN-GROUP WITH HIGH LEVEL OF PREJUDICE; BOTTOM RIGHT: IN-GROUP WITH LOW LEVEL OF PREJUDICE)

Study 2 - Rating the generated composites

Method

Participants and design

Participants. In total, 116 persons participated in this online study. Twelve were Dutch (3 male, 9 female), 96 were German (44 male, 52 female) and 7 had another nationality (3 male, 4 female). Age varied from 17 to 57 (M = 25.94, SD = 8.92). Participants were approached by social network websites such as Facebook and by email. This left us with a convenience sample, because only those with a Facebook- or Email account could be reached. 198 of the approached participants started the survey, but only 116 completed it. Consequently, these 82 uncompleted surveys were omitted from further analysis.

Design. The design of the study was a 1-factor between participants design, with group membership as a between-subjects variable. Participants were randomly assigned to one of the following two conditions: either the in-group target condition (n= 59), in which participants were exposed to a Caucasian perpetrator, or the out-group target condition (n= 57), in which participants were exposed to a Moroccan perpetrator. It took the participants approximately 15 minutes to complete the survey.

Procedure

As this survey was an online study, we were not able to have any influence on the location where participants completed the study. At the beginning of the survey participants were given a short introduction and were asked to give their informed consent. The survey consisted of three parts. In the first part, participants were told that there had been a robbery at the local night store and that the police, with the support of several eyewitnesses, had been able to compose a sketch of one of the offenders (Figure 1). Participants were asked to take a close look at the

sketch in order to memorize it as good as possible. Furthermore they were told that noise was added to the sketch in order to make it more difficult. Subsequently they saw 6 faces in a line up from which they had to choose the possible offender. At the beginning of the second part participants were presented with the photo of one of the possible offenders in the line-up. We told them that the experiment randomly selected a face, but we actually always used the photo of the true offender. They had the opportunity to take a close look by scrolling back and forth between different viewing angles of the photo (Figure 4). It lasted at least 20 seconds until they could press the "continue" button to get to the next page. Subsequently they were shown the composite sketches of the offender that were constructed in the first study (example in Figure 5). For each of these sketches, we asked participants to indicate the resemblance between the sketch and the offender they just saw. In the last part participants had to rate the outer impression of each of the sketches. They were provided with 10 adjectives for each sketch in order to rate how aggressive the sketch appeared to the participants. At the end of the survey participants had to provide their demographics and indicate to what extent they were in touch with Caucasian or Moroccan persons.

FIGURE 3: BASE FACE: MORPH OF CAUCASIAN AND MOROCCAN PERPETRATOR

FIGURE 4: EXPOUSRE TO PARTICIPANTS: TARGET PHOTO OF CAUCASIAN AND MOROCCAN PERPETRATOR

FIGURE 5: EXAMPLES OF PRODUCED SKETCHES OF CAUCASIAN AND MOROCCAN PERPETRATOR

Materials

The twenty-one participants of study 1 provided us with 21 classification images: 10 classification images for the Moroccan offender and 11 classification images for the Dutch offender. Furthermore we constructed one final classification image for each condition by combining the several classification images into one average classification image. We used the online research tool Qualtrics to construct and spread our survey.

Dependent on the condition the participant was assigned to, he or she had to rate 11 (12) classification images, thus the individual classification images as well as the final morph for each condition. In order to judge the resemblance between the actual perpetrator and the classification images, participants were confronted with the question "To what extent does the sketch resemble the offender?". They could respond on a 5-point Likert Scale from 1 (*not at all*) to 5 (*very much*). To get a clear picture of the appearance of the offender, we asked the participants to indicate to what extent they thought the sketch embodied several characteristics (see Appendix for all stimuli). We provided participants with five positively associated (i.e., *kind*, *charming* or *friendly*) and five negatively associated adjectives (i.e., *aggressive*, *violent* or *assertive*). Here as

well participants could respond on a 5-point Likert Scale from 1 (*not at all*) to 5 (*very much*). Cronbach's alpha for this scale was 0.65.

Results

Psychological valence of the faces. The second study used a two-factorial design with level of prejudice as within-subject variable and ethnical background as between-subject variable. It measured the psychological valence of the faces and was completely completed by 116 participants; most of them were students of the University of Twente in Enschede, the Netherlands. Eighty-two incomplete or incorrect trials were excluded from the analysis. Faces of in-group perpetrators produced by highly prejudiced participants were rated as most aggressive (M = 1.07, SD = 2.28), followed by faces of out-group perpetrators composed by highly prejudiced individuals (M = 1.04, SD = 3.07). The pictures of out-group perpetrators created by participants with a low IAT score were rated as more negative looking (M = -2.26, SD = 3.92) than faces of out-group perpetrators produced by individuals with a low IAT score (M = 0.01, SD = 1.96). On average, the pictures composed by low prejudiced participants were rated as less negative (M = 0.77, SD = 3.82) than those composed by highly prejudiced individuals (M = 2.12, SD = 3.52). To test whether this effect was significant and only true for the out-group condition as predicted in the hypothesis, a general linear model repeated measures ANOVA was conducted. Faces which were produced by respondents with a high level of prejudice looked more negative to participants in study 2, Wilk's Lambda= 0.87, $F(1, 114) = 17.64$, $p < .001$. However, no main effect for ethnic background ($F(1, 114) = 0.78$, $p = .38$) and no interaction effect between those two factors was found ($F(1, 114) = 4.2$; $p = .43$). Thus, the ethnic background of the perpetrator had no significant effect on the produced pictures in terms of psychological valence. Individuals

scoring high on the IAT, produced for both the Dutch (in-group) and Moroccan (out-group) condition more negative looking pictures.

Discussion

Eyewitness identifications are of crucial importance for judicial investigations. Therefore, the poor results present composite systems produce, are insufficient. Misidentifications by eyewitnesses are by far the greatest cause for wrongful convictions. Also, the ethnic background of the defendant seems to play a role. Most of the wrongfully convicted persons belong to minority groups (Innocence Project, n.d.). Therefore, this study investigated to what extent the level of prejudice of an eyewitness predicted the perceived psychological valence of a face of an either Dutch (in-group) or Moroccan (out-group) perpetrator.

In general, the hypothesis that more prejudiced eyewitnesses would produce more negative looking faces of out-group perpetrators was not confirmed. The Moroccans are a highly stigmatized group in the Netherlands and the Dutch majority labels them with adjectives like criminal or aggressive (Gordijn, Koomen, & Stapel, 2001). In a study by Dotsch et al (2011), highly prejudiced participants overallocated Moroccan faces to the label “criminal”. Therefore, it was expected that more prejudiced individuals would create pictures which appear more negative looking than composites produced by less prejudiced participants. The results of the present study showed that more prejudiced participants constructed more negative looking pictures for both the Dutch and Moroccan condition. Thus, the effect of prejudice was established for the out-group condition as well as for the in-group condition.

In our literature review we did not find a study which compared the effect of prejudice between in-group and out-group perpetrators. Dotsch et al (2011) did not made use of an in-group

condition. Therefore, this study is the first one to show that the effect of prejudice is not only observable for faces of out-group members but also for out-group members which counteracts the argumentation of earlier studies. Thus, the argumentation of Dotsch et al (2011) cannot be supported by these findings. They stated that the over-allocation of out-group faces of Moroccan perpetrators to the label criminal was a consequence of the negative stereotypes participants hold against them. Our findings suggest that the ethnic background of the face had no influence on the evaluation of the faces. Only the level of prejudice was relevant. At first glance, those results contradict each other. If prejudice is addressed towards out-group members, why should their ethnic background not be influential? A possible explanation would be that highly prejudiced individuals hold more negative stereotypes against criminals in general. Thus, the race of the perpetrator would be less influential than the fact, that he possibly convicted a crime. This hypothesis is supported by a study which showed that the higher support of the death penalty in the United States by Whites was mediated by their level of prejudice (Barkan & Cohn, 1994; (Barkan & Cohn, 2000). Other research showed that stereotypes vary in their content and strength: "The strength of a stereotype is the *degree* to which people believe the content of a stereotype characterizes a group of people" (Madon, 1997). Thus, one stereotype may have another content and may be activated faster than another one. In the present study, the stereotype about criminals may be stronger than the stereotype about Moroccans. Combining the results of Barkan and Cohn (1994; 2000) and Madon (1997) could explain the findings of the present study: participants with high level of prejudice have also stereotypes about criminals. Confronting them with faces of Moroccan criminals activate both stereotypes about the Moroccan and the criminal out-group. Their stereotype about criminals may have greater strength and therefore exceeded the stereotype about Moroccans. As a consequence, they produced more negative looking pictures for both the in-group and the out-group perpetrators, because both

groups consisted of (potential) criminals. This could also account for the discrepancy between the findings of Dotsch et al (2011) and our study.

In contrast to the study conducted by Dotsch et al (2011) the present study made use of individual faces of perpetrators. On this basis the participants created the composites. Dotsch et al (2011) asked the participants to think of what a typical Moroccan face looks like and to construct the composite on that basis. Thus, their stereotype was activated on a group-level while this study used an individual approach. Thereby, participants were triggered to think of their stereotype of Moroccans. It is questionable if Dotsch et al (2011) would have found a difference in the effect of stereotype if they had used an in-group condition as well. If this would be the case, the activation on group-level could be a reason for the difference between the in-group and out-group condition. This should be investigated in further research.

Previous research has shown that morphing produces better results for the identification of suspects (Bruce et al, 2002). The present used the reverse-correlation image classification technique to create morphs of in-group and out-group perpetrators. It was expected, that out-group morphs created by highly prejudiced individuals would be rated as more negative looking. The results showed that both the in-group and out-group morphs of participants with a high level of prejudice had a more negative perceived valence. The reverse-correlation image classification technique is thus also prone to bias resulting out of prejudice of the composite creators, but still a promising method for the production of composites. It takes previous findings into account since it based on recognition, follows a holistic face representation approach and uses the method of morphing. All those factors have been shown to contribute to a better functioning system for the identification of suspects (Davies & Donald, 1982; Tanaka & Farah, 2003; Bruce et al, 2002).

This discrepancy between our expectations and our results may also be due to some limitations of the present study.

Limitations of the present study

Our sample is not representative for the Dutch population since it was a convenience sample. As a consequence, selection biases can occur, which clearly was the case in the present study.

At first, most of the participants were students of the University of Twente. Thus, they do not represent the whole population in terms of age, level of education, contact with minority groups and other parameters. Besides that, most of them were German. Our study used sketches for the in-group condition of Dutch perpetrators. As a consequence, it can be assumed that German did not recognize them as real in-group members because of their nationality or may even have negative feelings towards Dutch faces which would have a negative influence on the validity of the IAT. Besides that, Moroccans are a well-known minority group in the Netherlands, but this is not true for Germany where Moroccans are less represented (Statistisches Bundesamt, 2013). Thus, German participants may not see the Moroccans as a minority group which also would have a harming impact on the IAT. A second shortcoming of the sample composition is that most of the participants studied psychology. Psychology students may be biased in that way that they tend to scrutinize every study in which they participate thereby possibly distorting their results. This may be especially true for this study because it makes use of the IAT which is a familiar instrument for psychology students. Since it measures the implicit prejudice of individuals, familiarity with the test clearly harms its validity.

Some students provided us with feedback, both in oral or textual form. For study 1, most participants found it hard to concentrate for such a long time (about one hour) only selecting between black-and-white faces with noise on them. Before beginning the study, we advised them to take breaks but no one made use of this offer. In the next study, it may be advisable to set fixed

breaks as a part of the study to ensure that participants can maintain a reasonable level of concentration or to reduce the number of trials.

This is also true for the second study. We received feedback which said that it was not clear how long the study took. Therefore, it may be helpful to add a bar to the bottom of the page which shows the progress of the study. Unfortunately, this element was missing in the present study which demotivated some participants to fill in the online questionnaire in a serious manner.

Of course, an online study in front of a PC does not resemble the high-stake situations eyewitnesses may be confronted with in a real judicial setting. They may be nervous or even frightened which impacts their performance as eyewitness. Therefore, this study can only be seen as an approach to investigate the role of prejudice in such a setting.

Suggestions for further research

To gain better insight into the relationship of level prejudice of an eyewitness and his or her produced composites further research should try to rule out the above mentioned shortcomings of this study. Beside some minor changes in the appearance of the study, it is important to test a representative sample of the target population.

There is an enormous amount of research on both the cross-race effect and applied research on eyewitness composite systems (Allport & Kramer, 1946; Anthony et al 1992; Bernstein et al, 2007 Bruce et al, 2002; Davies & Donald, 1982; Dotsch et al 2011). But the link between those subjects is still insufficiently investigated. Since this relationship seems to be of crucial relevance in a judicial setting, it has to be further examined. Composite systems have to be improved to take the possible role of prejudice and social categorization and its effects on the produced pictures into account.

Especially, the difference between individual activation and activation on group-level of the stereotype has to be further investigated. Therefore, a combination of the study of Dotsch et al

(2011) and the present study could be a possible approach in which an in-group condition is added to the makeup of Dotsch et al.

Studies investigating real life settings may be of great importance since an only study barely resembles the situation eyewitnesses may be confronted with. Feelings of great responsibility or even fear may interfere with their performance as an eyewitness. Therefore, real cases should be studied.

Conclusion

The present paper used the reverse-correlation image classification technique to test whether prejudice has an influence on the creation of composites. Results showed that highly prejudiced participants produced more negative looking pictures for both the in-group and the out-group perpetrators' faces. It hast to be further investigated, which mechanisms play a role during the activation of stereotypes in a judicial setting. Especially, the difference between stereotype activation on group-level and individual level should be examined.

References

Allport, G. (1979). *The Nature of Prejudice.* Perseus Books Publishing, US: New York.

Allport, G., & Kramer B.M. (1946). Some root of prejudice. *Journal of Psychology, 22,* 9-39.

Anthony, T., Copper, C., Mullen, B. (1992). Cross-racial facial identification: A social cognitive integration. *Personality and Social Psychology Bulletin, 18,* 296-301.

Barkan, S.E., & Cohn, S.F. (1994). Racial Prejudice and Support for the Death Penalty by Whites. *Journal of Research in Crime and Delinquency, 31*(2), 202-209.

Barkan, S.E., & Cohn, S.F. (2000). Why Whites Favor Spending More Money to Fight Crime: The Role of Racial Prejudice. *Social Problems, 52*(2), 300-314.

Blech, Jörg (2007, December 3). Die Unschuldigen befreien. *DER SPIEGEL.* 144-146.

Bernstein, M.J., Young, S.G., Hugenberg, K. (2007).The cross-category effect: mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*(8), 706-712.

Bruce, V., Ness, H., Hancock, P.J.B., Newman, C., & Rarity, J. (2002). Four heads are better than one: combining face composites yields improvements in face likeness. *Journal of Applied Psychology*, 87, 894.

Capozza, D., Boccato, G., Andrighetto, L., Falvo, R. (2009). Categorization of Ambiguous Human/Ape Faces: Protection of Ingroup but Not Outgroup Humanity. *Group Processes Intergroup Relations, 12*(6), 777-787.

Castano, E., Yzerbyt, V., Bourguignon, D., & Seron, E. (2002). Who may enter ? The impact of in-group identification on in-group/ out-group categorization. *Journal of Experimental Social Psychology, 38,* 977-997.

Chance, J.E., & Goldstein, A.G. (1981). Depth of processing in response of own- and other-race faces. *Personality and Social Psychology Bulletin, 7*, 475-480.

Coenders, M., Lubbers, M., Scheepers, P., Verkuyten, M. (2008). More than Two Decades of Changing Ethnic Attitudes in the Netherlands. *Journal of Social Issues, 64*(2), 169-185.

Cross, J.F., Cross, J., Daly, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. *Perception & Psychophysics, 10*, 1939-396.

Davies, G., & Donald, C. (1982). Face recall. An examination of some factors limiting composite production. *Journal of Applied Psychology, 67*(1), 103-109.

Davies, G., & Oldman, H. (1999). The impact of character attribution on composite production: A real world effect? *Current Psychology: Developmentak, Learning, Personality, Social, 18*, 128-139

Davies, G.M., & Valentine, T. (2006). *Handbook of eyewitness psychology* (2nd ed.) Mahwah, NJ: Erlbaum.

Dotsch, R., Wigboldus, D.H.J., van Knippenberg, A. (2011). Biased allocation of faces to social categories. *Journal of Applied Psychology, 100*(6). 999-1014.

Elliot, D.N., & Wittenberg, B.H. (1955). Accuracy of identification of Jewish and non-Jewish photographs. *Journal of Abnormal and Social Psychology, 51*, 339-341.

Ellis, H., Davies, G. M., & Shepherd, J. (1978). A critical examination of the photofit system for recalling faces. *Ergonomics*, *21*(4), 297-307.

Eysneck, M., & Kaene, M.T. (2010). *Cognitive Psychology: A Student's Handbook* (6th ed.), Florence, Kentucky: Psychology Press.

Frijda, N.H., The Laws of Emotion. *Human Emotions: A Reader*. Malden, MA: Blackwell Publishers.

Gordijn, E.H., Koomen, W. & Stapel, D.A. (2001). Level of prejudice in relation to knowledge of cultural stereotypes. *Journal of Experimental Social Psychology, 37,* 150-157.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.K.L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

Harris, L.T., & Fiske, S. (2006). Dehumanizing the Lowest of the Low. *Psychological Science, 17*(10), 847-853.

Hasel, L.E., & Gary, L. (2007). Facial Composite Production by Eyewitnesses. *Current Directions in Psychological Sciences, 16*(1), 6-10.

Himmelfarb, S. (1966). Studies in the perception of ethnic group members: I. Accuracy, response bias, and anti-Semitism. *Journal of Personality and Social Psychology. 74*, 1646-1480.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. Journal of Experimental Psychology: General, 128, 3–31.

Innocence Project (n.d.). The causes of Wrongful Conviction. Retrieved March, 23, 2015, from http://www.innocenceproject.org/causes-wrongful-conviction

Kosic, A., & Phalet, K. (2006). Ethnic categorization of immigrants: The role of prejudice, perceived acculturation strategies and group size. *International Journal of Intercultural Relations, 30,* 769-782.

Madon, S. (1997). What do people believe about gay males? A study of stereotype content and strengths. *Sex Roles, 37*(9-10), 663-685.

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

Mangini, M.C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science*, 28, 209-226.

Quanty, N.B., Keats, J.A., & Harkins, S.G. (1975). Prejudice and Criteria for identification of ethnic photographs. *Journal of Personality and Social Psychology, 32*, 449-454.

Sangrigoli, S., Pallier, C., Argenti, A.M., Ventureyra, V.A.G., de Schonen, S. (2005). Reversibility of the other-race effect in face recognition during childhood. *Psychological Science, 16*, 440-444.

Shephered, J.W., Ellis, H.D., McMurran, M., & Davies, G.M. (1978). Effect of character attribution on Photofit construction of a face. *European Journal of Social Psychology, 8*, 263-268.

Simon, D., Pham, L. B., Le, Q. A., & Holyoak, K, J. The emergence of coherence over the course of decision making. Journal of Experimental Psychology: Learning, Memory, and Cognition, 27, 1250–1260.

Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. Journal of Personality and Social Psychology, 86, 814–837.

Statistisches Bundesamt. (2013). *Bevölkerung und Erwerbstätigkeit: Bevölkerung mit Migrationshintergrund - Ergebnisse des Mikrozensus 2013.* Retrieved from https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/MigrationIntegration/Migrationshintergrund2010220137004.pdf?__blob=publicationFile

Tanaka J. W., & Farah M. J. (2003). *The holistic representation of faces in Perception of Faces, Objects, and Scenes: Analytic and Holistic Processes*. New York: Oxford University Press.

Taubert, J., Apthorp, D., Aagten-Murphy, D., & Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision Research,51*(11), 1273-1278.

Todorov, A. & Dotsch, R. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, *3*, 562-571.

Verkuyten, M., & Zaremba, L. (2005). Interethic relations in a changing political context. *Social Psychology Quartely, 68*, 375-386.

Weiss, D.J., & Yoes, M.E. (1991). *Advances in Educational and Psychological Testing. New York: Springer.*

Wells, G.L., & Hasel, L.E. (2007). Facial Composite Production by Eyewitnesses. *Current Directions in Psychological Science, 16*(1), 6-10.

Wittenbrink, B., Judd, C.M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262-274.