Cognitive Discrimination

Michèle Belot*

March 10, 2010

Abstract

This study provides a first set of experimental results highlighting a new mechanism for racial discrimination, based on a cognitive limitation in facial re-identification across races. An experiment is conducted to study how people record and recall payoff-relevant information about others, within and across races. East-Asian and White subjects see an equal number of pictures of East-Asian and White faces and each face is mapped to a payoff-relevant value. Incentives are provided to recall faces associated with higher values. We observe a clear asymmetry in the accuracy of recall: High value faces are more accurately recalled within race than across races. I contrast these results with a treatment where race is a distinctive attribute for a minority of faces (where there is only a small number of pictures of East-Asians). In that case, I find that East-Asians are favored both by White and East-Asian subjects. That is, the biases identified in the first treatment cannot be attributed to prejudice or preferences. These results raise new questions on the implications of such cognitive biases for the nature of cross-racial relations, in particular for phenomena relying on reidentification, such as the formation and maintenance of social ties, the establishment of trust and the sustainability of cooperation.

Keywords: Own-Race-Bias, Discrimination, Face Recognition, Bounded Memory, Identity

1 Introduction

Remembering people plays a crucial role in many social contexts, for example in the formation and maintenance of social ties, in repeated social interactions and in mechanisms relying on repeated interactions such as trust and cooperation.

*The author would like to thank Pavan Mamidi, Diego Gambetta, Philipp Kircher, Marco van de Leij, Uri Gneezy, Scott Blinder, Luis Miller, Scott Moser and audiences at various seminars for useful comments. The author is also extremely grateful for the programming support offered by her father, Jacques Belot.

In a world where individuals differ in their intentions or values for economic transactions and where information about these values or intentions are not directly observable or contractible, keeping track of who is doing what helps to coordinate decisions and efficiency. It enables people to choose whom they interact with, and condition their behavior on the information they possess about other people. For example, those who cooperated in the past can be rewarded and those who defected, punished.

Facial recognition is one of the key technologies used for individual re-identification. In many social interactions, a face is all we have or we rely upon to re-identify someone. Local merchants typically know their customers only by face, not by name or any other identifier. Business environments are structured around social events where facial recognition plays a key role. In fact, asking for the name of a person one has met before is usually experienced as embarrassing. Strikingly, there appear to be large differences in the ability to remember and recognise faces between and within races. A robust empirical finding in the literature in psychology is indeed the presence of an "own-race bias" in face recognition - the fact that people are generally better at rembering faces from their own race than other races (see Meissner & Brigham, 2001; Slone, Brigham, & Meissner, 2000 for reviews). This cognitive bias seems prevalent among all racial groups (Ng & Lindsay, 1994; Teitelbaum & Geiselman, 1997), although the evidence suggests the effect is most pronounced for Caucasians viewing members of racial minority groups (Meissner & Brigham, 2001). Psychologists have repeatedly stressed the implications of such bias for eyewitness identification in the criminal justice system. Eyewitness misidentification appears to be *the* single greatest cause of wrongful convictions in the US, playing a role in more than 75% of convictions overturned through DNA testing.[1] While the implications are particularly dramatic in the context of crime, such bias in recall may have ramifications in many social contexts: from whom we become friends with, to whom we trust and whom we recommend for a job, etc. Biases in re-identification across race may shape the structure of cross-racial relations, and possibly have discriminatory implications.

Yet, in a social context, the critical aspect is not face recognition itself, what matters is the retrieval of *information* associated to the person or face, infor-

[1] Source: "The Innoncence Project"; a national litigation and public policy organization dedicated to exonerating wrongfully convicted people through DNA testing (www.innoncenceproject.com)

mation relevant to the social interaction. A *reputation* is precisely this mapping between identities and information. For example, in a trust situation, keeping track of players'identities is useless unless these identities are mapped to information about, for example, the players' past actions (e.g. who cooperated and who defected). To understand the *economic* relevance of possible cognitive limitations in re-identification, it is necessary to study the *joint* recall of identities and payoff-relevant information associated to these identities. The question is whether there are biases arising in the retrieval of information associated with a person, for example, whether a teacher is less likely to remember a good student if he is from another race than her own. The goal of this study is to provide a first set of experimental results on this specific process of joint recall of identity and payoff-relevant information. The findings in the literature in psychology are not directly informative on the matter because we do not know whether better facial recognition implies better retrieval of information associated with people and more efficient decisions. Hanley (2008) documents the many situations in which people report being able to remember faces but cannot recall the context, that is the recall of information - and in particular payoff-relevant information - may not necessarily exhibit the same biases. Particularly because differences in payoffs may affect the way information is recorded in the first place. People may put more or less effort into remembering people, depending on their "value" for social interaction.

We think of three classes of applications where the phenomenon is likely to be relevant. A first class of applications relates to the literature on network formation. Any social tie starts with a first interaction and then requires re-identification. It is well-known that social networks are heavily biased towards one's own race - a phenonomenon known as *racial homogamy* (McPherson et al. (2001), Fong and Isajiw (2000)).The literature in sociology and economics has mainly focused on two possible explanations for this phenomenon: one is *homophily* (relationships within race are valued more than across race), the other is the frequency of meeting or *opportunities* (people may be more likely to meet and interact with people of similar race) - see Currarini et al. (2009) for a model of network homogamy taking these two forces into account.[2] Surprisingly, the mechanism of re-identification - and the possible racial biases associated with

[2] Interestingly, the homophilic feature of networks seems to appear very early on. Mollica et al. (2003) study friendship ties among MBA newcomers and find that friendship networks are already segregated 6 weeks after the beginning of the academic year.

it - despite being an essential mechanism in the formation and maintainance of social ties has not been considered at all.

A second class of applications regards mechanisms relying on the repeated character of interactions and re-identification, such as trust and reciprocal altruism (Axelrod, 1981), enhancing efficieny in environments where transactions cannot easily be governed by contracts. These mechanisms are often very demanding in the information structure required to sustain cooperation - in terms of player's identities and records of past behaviour. Basu et al. (2009) show that facilitating recordkeeping with an external support (e.g. allowing subjects to take notes) enhances cooperation. Cognitive limitations in cross-race re-identification may therefore play a role in hampering cooperation across races and possibly compromise the use of these mechanisms to enforce cooperative behavior across races.

A third class of applications is the practice of recommendation, which plays a large role in school and employment selection systems. The issue of recall of information beyond recorded information - typically the main added value of a recommendation - is likely to have salient implications in such settings. University lecturers are very familiar with the difficulties of writing a recommendation letter for students from large classes. As proposed by Fryer and Jackson (2008), cognitive biases in memory could have implications for hiring and placement decisions and constitute an alternative mechanism for racial discrimination.

These three examples of applications have been extensively studied in economics, together with their racial biases. I propose here a novel angle and argue that there could be a cognitive mechanism playing a key role in these phenomena and explain possible racial biases associated with them. The ambition of this study is to provide a first benchmark for the study of these cognitive limitations by conducting a simple and fundamental experiment.

The experiment proceeds as follows. Subjects see a sequence of pictures of faces from East-Asian and White people, and each face is mapped to a randomly allocated value, that appears simultaneously on the screen. In a second stage, subjects are asked to select a number of pictures to enter a lottery - the lottery picks one picture at random that will determine the subject's final earnings. At the stage of selection, subjects see the same set of pictures again - but without their associated value. I investigate the accuracy in joint recall of identity and information, more specifically, I evaluate the optimality of selection decisions and compare the efficiency of decisions within and across race. The experiment has a neutral frame and the values could be interpreted in different ways. They are a general measure of payoff-relevance, and arise in many models of social interactions: In repeated games the value of a person represents the benefit from the interaction (which depends for example on the number of times a person defected in the past), in network formation games they represent the value from linking with a person, and in labor markets they could simply represent a measure of the productivity of the person.

There are two major advantages of a controlled environment in this context. First, we can strictly limit the consequences of decisions to the decision-maker: the selection decision only affects the subject's payoff and the person who is the object of the decision (the face on the picture) does not incur any loss or a gain by being selected or not. This rules out a role for other-regarding considerations, such as fairness or willingness to provide benefits to own group members over others. Second, the individual payoffs associated with each face are fully controlled for; and subjects are fully informed about these individual payoffs and about the procedure of assignment of payoffs to faces. This exogenous assignment means that the faces alone do not contain any information about these payoffs. This is important given the recent evidence showing relationships between facial features (and in particular race) and inferences about personal characteristics such as competence, trustworthiness, etc. (Todorov et al. (2005), Eckel and Petrie (2008), Rule and Ambadie (2008), Todorov and Duchaine (2008), Duarte et al. (2009)). Here the only cognitive mechanism enabling an inference from faces to values is memory. Thus, the environment aims at isolating the role of recall in decisions and neutralise the role of all other factors possibly affecting decisions in real life, in particular preferences, perceptions or stereotypes. However, one might still worry that these alternative mechanisms are not entirely neutralised by the experimental design. To verify that memory is indeed the core mechanism at work in these decisions, I exploit an interesting dual aspect to cross-race re-identification. Race appears to be a distinctive attribute - it has been found to be a prime characteristic encoded about others (Montepare and Opeyo (2002)) - and distinctiveness has been found to enhance re-identification (Shepherd et al. (1991), Tibbetts and Dale (2007)). Thus, if there are few people of a given race, it should enhance their re-identification (in a crowd of White people, an East-Asian man will stand out). If memory drives decisions, then decisions involving those with a distinctive attribute should be more efficient. To shed light on this hypothesis, I compare decisions across two treatments, one where there are as many East-Asian pictures as White pictures and another where there are only two East-Asian pictures, a man and a woman.

The data show some striking patterns: First, there is a clear own-race bias in re-identification when there is an equal number of faces of each race. This bias is most pronounced among white participants, who recall much more accurately high value faces from their own race than other races (68% of the white faces with a value among the eight highest values are selected against 50% for East-Asian faces). Second, an equal number of faces of each race enters the selection. Thus, we find evidence of negative racial discrimination at the top of the value distribution (high value faces of the other race are less likely to be selected). Lower ranked faces are more likely to enter the selection across races than within race. Finally, we find that when there is a minority of pictures of East-Asians, then decisions involving East-Asian pictures are significantly more efficient for both White and East-Asian subjects. The own-race bias turns into a minority bias.

The rest of the paper is structured as follows. Section 2 discusses the relevant literature in psychology and economics. Section 3 presents the experimental design and Section 4 the results. Finally, I conclude in Section 5 and discuss new research questions.

2 Related literature

2.1 The Own-Race Bias

Psychologists and neuropsychologists have studied extensively the cognitive and neurological processes involved in facial recognition (see Duchaine (2008) for a review). The own-race bias (ORB) is one of the most robust empirical findings in the literature on facial recognition. Meissner and Brigham (2001) provide a detailed meta-study of the last thirty years of literature, reviewing 39 articles involving the responses of over 5,000 subjects. There are a few studies that fail to find a cross-race effect. The overwhelming consensus among social psychologists is that these effects not only exist, but are quite large.

The cognitive and social factors responsible for the ORB remain unclear

(Slone et al., 2000). Theories proposing that the degree of interracial contact should be negatively associated with level of ORB have been only weakly supported (Chiroro & Valentine, 1996). Meissner & Brigham (2001) show in their meta-analysis that interracial contact accounts for only about 2% of the variance in ORB across samples. Although negative racial attitudes are correlated with limited interracial contact, no relationship has been found between the ORB and racial attitudes, whether explicit or implicit (Ferguson, Rhodes, & Lee, 2001). Training does not seem to help much either. Lavrakas et al. (1976) show that training could reduce the magnitude of the ORB, but the effect was short-lived: One week later there was no difference between trained and untrained participants.

Some evidence suggests that one reason for the ORB may be that cross-race faces are processed differently than own-race faces. In essence, cross-race faces may be perceived more "holistically" - more like objects - (Rhodes et al., 1989; Tanaka et al., 2004). This idea is confirmed by neurophysiological studies. Neurophysiologists have identified specifc areas of the brain active in the processing of faces and that the processing of cross-race faces is different from the processing of own-race faces (Golby et al. (2001), Cunningham et al. (2004) and Duchaine (2008) for a recent survey).

Levin (1996, 2000) proposes that other race effects are caused by selection of different facial features in same and other race faces. Whereas individuating information is selected in same race faces, race specifying information is emphasized in representations of other race faces at the expense of individuating information. In fact, race has been shown to be one of the prime characteristics encoded in human interactions, together with gender and age (Montepare and Opeyo (2002)). On the other hand, a number of studies show that faces rated as distinctive are more accurately remembered (Shepherd et al. (1991) and Valentine, 1992 for a survey of relevant studies). Thus in a situation where race is a scarce attribute, it could serve as an abious marker of identity and improve recognition significantly.

Recent work suggests that face recognition develops with age. Pascalis O. et al. (2002) showed that 6-month old infants, 9-month old infants and adults were able to discriminate between human faces but only 6-month-olds could discriminate between monkey faces. This phenomenon is similar to the loss of sensitivity to phonemes not used in the infant's native language (Werker and

Tees (1984), Kuhl (1992), Aislin et al (1998)). Differential processing of faces of different races follows a similar developmental course.

2.2 Joint recall of identities and payoff-relevant information

As mentioned earlier, to understand the economic implications of cognitive biases in re-identification, one needs to study the joint recall of identities and payoff-relevant information. The most related studies in psychology are those that study the recall of *associations* between faces and information. The seminal work in that area is Taylor et al. (1978), who studies how subjects recall the contents of interactions between people of mixed gender and race ("*Who said What?*")[3]. They show that subjects are more likely to misattribute statements of people of the same race than different races. The conclusion from these studies is that race and gender are primal attributes encoded by the brain. In turn, racial categorisation has been seen as underpinning the emergence of racial stereotypes.[4]

From the point of view of an economist, the question is how do people record and retrieve an information of the type "Individual i has productivity y". Payoff relevance implies that the information has an ordinal or even possibly cardinal nature - payoffs can be ordered and differences in payoffs can be more or less pronounced. Fryer and Jackson (2008)[5] propose a model of how such information might be stored in memory. They also conjecture that people are sorted into categories, and these categories may also correspond to payoffs. Each category has a prototype - a unique vector of attributes. The main idea is that people keep track of the variation in attributes *across* categories but not within a category. People sorted in the same category are blended together. For example, one could sort people into two categories; high productive and low productive people. They argue that minority groups may be sorted into coarser categories than majority members because they are less likely to be involved in

[3] I am very grateful to Oliver Curry for informing me about this literature

[4] Interestingly, the view that race encoding is wired into our brain has been challenged recently. Kurzban et al. (2001) argue that racial categorisation can be eliminated if other social coalitions are more relevant (for example, whether they belong to a particular social group). They conjecture that racial categorisation is a by-product of an evolutionary advantageous behaviour consisting in sorting people into "in-group" and "out-group" members.

[5] Their model fits in the growing literature on the role of cognitive limitations in memory and their implications for social interactions (Mullhainathan (2002) and Mullhainathan et al. (2008)).

frequent repeated interactions. This will result in asymmetric treatment even if the two populations are identical in terms of productivity distribution, and even if there is no malevolent taste against minority groups.

To our knowledge, there is no experimental evidence documenting the recall of identities and payoff-relevant information. We know very little about how such cognitive biases may affect the efficiency of decisions, how categorisation may operate in the presence of payoff-relevant information.

The economic implications of a cognitive racial bias in person recognition are similar in nature to those present in contexts of asymmetric information and noisy signals. The model of *screening discrimination* developed by Aigner and Cain (1977) is based on the idea that employers may be better able to evaluate the productivity of workers from their own race than other races. They derive the wage implications of differences in the variance of inferred productivities across groups. They show that groups with a higher variance (and therefore a more noisy signal) would have a more compressed wage distribution. Similarly here, a bias in the memory technology translates into noisier information. If the information cannot be retrieved, the expected payoff of a person will correspond to the average payoff corresponding to her observable identifiers (race in particular). Notably, this form of discrimination does not necessarily lead to a better or worse *average* treatment of a racial group. What it does is compressing the distribution of rewards - the better ones are less likely to be rewarded and the worst ones are less likely to be punished. In the example of eyewitness identification, the own-race-bias implies that a criminal is more likely to be wrongly convicted if he is from a different race, but it also means that the true criminal is more likely to walk away without being punished.

3 Experimental Design

The experiment was conducted at the laboratory of the Nuffield Centre for Experimental Social Siences (Oxford). 106 subjects were recruited, 60 Caucasians and 46 East Asians. Invitations were sent by e-mail to subjects in the pool with East Asian and British last names[6]. Subjects received a £4 show-up fee and an additional payment depending on the performance in the memory task (see explanation below). The instructions were read aloud and subjects could

[6] The invitation asked for participants between the ages of 18 and 30. We have excluded participants above 50 years old from the analysis (n = 3).

ask questions immediately thereafter. Subjects earned £8.6 on average and the experiment lasted for 30 minutes in total. A post-experimental questionnaire was handed in, asking for information about ethnicity, age, occupation, country of birth, age of arrival in the UK and a statement about their perceived re-identification ability.

The experiment is structured in 3 stages. The first stage is a *viewing stage*, participants view an automated sequence of 24 pictures, each picture is shown for 3 seconds and then the screen moves to the next picture.[7] Each picture appears on the screen together with its value. The second stage is a *selection stage* – participants see the same 24 faces again, but without their associated value, and are asked to select 8 pictures that will enter a lottery (in a third stage); a lottery that will determine their final earnings. Importantly, pictures are in a different sequence, and the pictures are from a different angle. The sequence is not automated at this stage and subjects can go back and forth between pictures for 3 minutes. In case of incomplete selection (less than 8 pictures selected by the subject), pictures are selected at random among the remaining non-selected pictures to complete the selection. Note that the sequence of presentation of faces was randomized as well for each participant, and for each stage. The third stage is a *lottery*, whereby one of the pictures from the selection in stage 2 is picked at random.[8] The subject's earnings are equal to the value of the picture divided by 10, in British Pounds.

Facial stimuli

Faces were drawn from a database provided by TARRLAB [9] All images were extracted from standard digital video (720x480), the background removed and the faces were scaled to be roughly equated in terms of size.The pictures were selected according to the following parameters: Race (East-Asian[10] and Caucasian White); shave/stuble no make-up; no beards or mustache; no facial hair or visible make-up; no glasses; natural hair (no wig); neutral affect; orientations:

[7] The choice of number of pictures, time and mix of gender is in line with the common practice in the psychology literature.

[8] Note that if pictures were added to complete the selection in stage 2, these enter the lottery as well. No matter how many pictures were selected in stage 2 by the subject, there were always 8 pictures entering the lottery. In practice, all participants selected 8 pictures exactly.

[9] Face-Place Face Database Project (http://www.face-place.org/); Copyright 2008, Michael J. Tarr. Funding provided by NSF award 0339122.

[10] East-Asian faces were selected in the Asian sample with the help of an Asian person.

0°, 15°left, 15°right, 30°left, 30°right. This left us with a database of 11 East Asian men, 15 East Asian women, 35 Caucasian men and 44 Caucasian women.

For each subject, a set of 24 faces was randomly chosen. Two pictures of each face, from a different angle, were ramdonly chosen (one used in the viewing stage and the other used in the selection stage). This is to make sure the exercise involves face recognition rather than picture recognition and prevents subjects from using other cues than the face itself to remember the person. The sequence of viewing is determined randomly for each participant, for both the viewing and selection stages.

A picture of a mixed race person has been chosen to illustrate the instructions (see instructions in the appendix).

The values of faces

A unique value was pre-assigned randomly to each picture. The values correspond to randow draws from a discrete normal distribution truncated at 10 and 70, with mean 40 and standard deviation 15. This range has been chosen so that all values have 2 digit numbers. The choice of the normal distribution is motivated by the relatively low presence of "outliers" in many real applications, outliers who are often those one would like to remember in taking decisions.

The instructions have been written carefully to inform subjects in detail about the procedure of assignment of values to pictures (see Appendix). In particular, the discrete normal distribution was represented graphically with a *roulette,* illustrating the differences of probabilities of occurence corresponding to each number from 10 to 70 by differences in areas. Subjects are also informed that the procedure is repeated for each picture independently. The instructions do not mention race at any point.

Treatments

Treatment 1 - "Equality treatment" - An equal number of East-Asian and Caucasian pictures are selected: 6 men and 6 women of each race.

Treatment 2 - "Majority-Minority treatment" The set is constituted of 22 Caucasian pictures (11 men and 11 women) and 2 East-Asian pictures (1 man and 1 woman).

4 Baseline results

4.1 Summary statistics

Tables 1 and 2 present summary statistics of the sample of subjects. It is worth pointing out in addition to these statistics that none of the East-Asian subjects was born in the UK and the average time spent in the UK is 1.15 years.

Table 1: Treatments and number of subjects

	Treatment Equality	Treatment Minority
Caucasian White subjects	32	29
East-Asian subjects	28	27

Table 2 - Subjects summary statistics

	Caucasian White subjects		East-Asian subjects	
	Equality	Minority	Equality	Minority
Age	25.3	24.7	25.6	25.3
	(5.4)	(6.2)	(4.4)	(5.4)
Share women	50%	50%	38%	41%
Mean score	47.6	50.3	48.2	45.4
	(5.4)	(12.1)	(4.6)	(11.5)

The baseline outcome of interest is the proportion of correctly allocated faces; that is, the proportion of faces selected among the faces with the 8 highest values The benchmarks are 33% for random selection and 100% for perfect memory.

Overall, subjects did significantly better than chance, with a proportion of correctly allocated faces equal to 61% ($P = .000$)[11]. The mean value of the selected pictures is 48 (S.E. .47), significantly higher than the benchmark for chance (40, $P = .000$).

4.2 Treatment *Equality*

Fig. 1 shows the average proportion of top 8 pictures selected, by racial groups of subjects and faces. The selection shows a substantial own-race bias[12]. The proportion of top 8 East-Asian faces selected by White participants is $q^{top8}_{ea,w} = .50$ against $q^{top8}_{w,w} = .66$ of top 8 White faces ($P = .001$). The opposite trend is

[11] All reported tests are two-sided Mann-Whitney tests with the individual proportions as independent observations.

[12] All reported tests correspond to post-estimation F-tests based on linear regression estimates pooling all choices made by the 104 participants, and clustering the standard errors by faces.

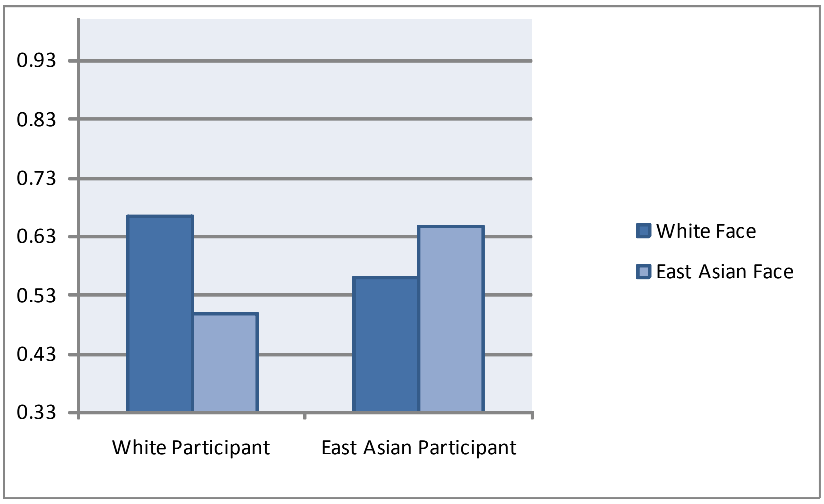

Figure 1: Proportion of top 8 faces selected - Treatment *Equality*

observed for East-Asian participants: A proportion $q^{top8}_{ea,ea} = .65$ of top 8 East-Asian faces is selected, against $q^{top8}_{w,ea} = .56$ of top 8 White faces), but the difference is not statistically significant ($P = .25$).

One immediate question is whether the inefficiencies in allocation decisions arise at the top of the distribution of the top 8 pictures, that is whether the biases affect those at the very top of the distribution or arise only for those who are ranked lower. To shed light on this matter, the same analysis is conducted for the top 3 pictures. The proportion of top 3 East-Asian faces selected by White subjects is $q^{top3}_{ea,w} = .60$, against $q^{top3}_{w,w} = .80$ for White faces($P = .036$). For East-Asian subjects, these proportions are $q^{top3}_{ea,ea} = .80$ for East-Asian faces and $q^{top3}_{ea,w} = .72$ for White faces, difference that is not statistically significant ($P = .41$). There are two important conclusions from these results. First, among the set of faces with a value qualifying for the top 8, those with a higher value are more likely to be included than those with a lower value, both within and across race. Values do increase the accuracy of recall. Second, this effect does not eliminate cross-racial biases, on the contrary. The difference in the probability of inclusion becomes even larger for faces included in the choice sets of White subjects.

Another way of looking at the data is to look at the frequency of selection

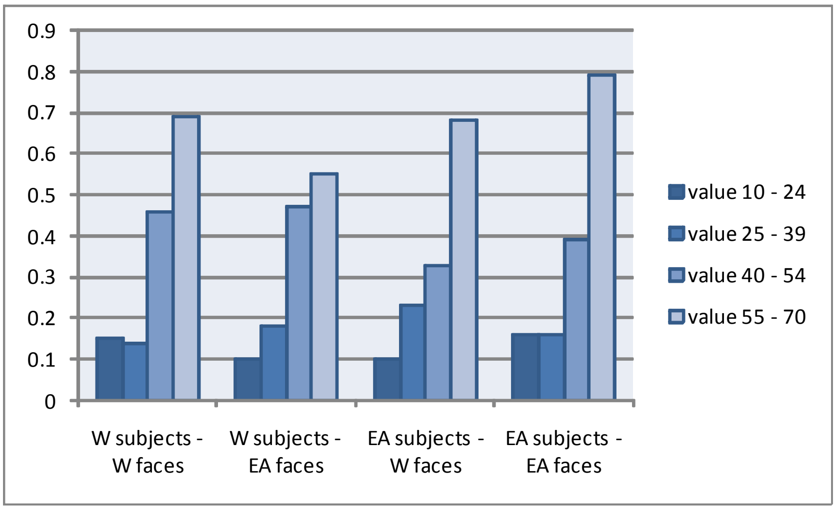

Figure 2: Proportion of selected faces as a function of value - Treatment *Equality*

as a function of the value rather than the rank. Indeed, at the stage of viewing, subjects do not know the actual distribution of values they will see. They know from the instructions that the values are drawn from a normal distribution with mean 40 and standard deviation 15. Possible rules of thumb could be to pay attention to faces with values above 40 or 55. Fig. 2 presents the probability of selection as a function of values according to the race of participants and faces. Again, the own-race bias appears clearly. High value faces are much more likely to be selected within race than across races, and this effect is stronger for white participants than East Asian participants. The difference in the probability of selection is mostly present at the top of the value distribution, for values of 55 and above.

In conclusion, the evidence shows an own-race bias when there is an equal number of faces of each race. The effect is mainly driven by differences in the probability of selection at the top of the value distribution and is more present among White participants. These effects are in line with the findings of the literature in psychology regarding facial recognition. It is worth pointing out though that East-Asian subjects are living in the UK, and therefore might be a selected sample with respect to the ability to distinguish white faces.

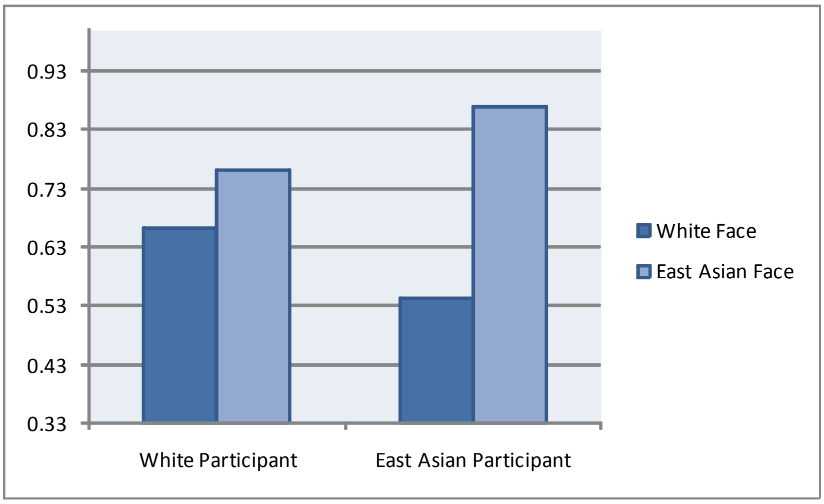

Figure 3: Proportion of top 8 faces selected - Treatment *Minority*

4.3 Treatment *Minority*

In the second treatment, the composition of facial stimuli only includes 2 East-Asian pictures, a man and a woman. The hypothesis tested here is whether race and gender can enhance identification when they are scarce attributes.

The results are supportive of the hypothesis (Fig. 3). While the overall misallocation hardly changes for East-Asian subjects, it greatly improves for white subjects; East-Asian faces are now more likely to be correctly allocated than in the first treatment and more so than white faces. The bias now takes the form of a *minorty bias*: faces from the minority group are more likely to be correctly allocated than faces from the majority group. For white subjects, the difference in the proportion of correctly allocated asian faces in comparison to white faces is .10 and for East-Asian subjects, this difference is .39. Both are positive, but the difference in these differences is also statistically significant $P = .007$), thus East-Asian subjects remain relatively better at allocating East-Asian faces relative to white faces than White subjects.

These effects are important to make sure that the biases identified in the previous section are not driven by other mechanisms - such as racial prejudice. If racial prejudice was at work and explained why high value pictures from the other race are less likely to be selected, then we should see this in both in treatments. The fact that high value East-Asians are now *more* likely to be selected by both groups means that prejudice cannot be the explanation for the biases found in the treatment *Equality.* Of course, this claim holds only if other racial biases do not vary with the relative sizes of racial groups. If they do - for example if selecting a member of the other race is more appealing when there are fewer of them, then it would seem that one could also explain this reversal of fortune across treatments. To address this concern, it is important to stress that the prediction corresponding to the treatment effect is not that people with a distinctive attribute should be *more likely to be selected,* the prediction is that the decisions involving them should be more *efficient*, that is, they should *increase* the probability that those with a *high value* are selected, but *decrease* the probability that those with a *low value* are selected. That is, the treatment effect should be to reduce confusion. Prejudice or other racial biases, on other hand, should lead to unbalances in the portfolio of pictures chosen. The exercise in the next section will be enlightening in that respect.

4.4 Composition effect or confusion?

The next question I ask is to what extent the inefficiencies are driven by differences in the probabibility of entering the selection (composition effect) or by confusion conditional on selection. I estimate the probability of entering the selection conditional on race and being in the top 8 and test whether faces of different races are more or less likely to enter the selection depending on whether they are in the choice set of White or East Asian subjects. Note that one implication of the chosen distributional form is that quite a large number of sets (18 out of 60 in treatment 1 and 8 out of 56 in treatment 2) contain more than 8 pictures qualifying for the top 8, that is, there are at least 2 faces corresponding to the eightst rank. Thus, by definition, not all top 8 pictures could possibly be selected. Table 3 presents two sets of regressions, one set including the whole sample and the other excluding choice sets with more than 8 pictures qualifyung for the top 8. The results are very similar in both cases and show very little biases in the selection process, for either racial group and in either treatment. Thus, the biases identified in the previous sections are not driven by differences in the probability of selection, but rather by differences in the values of those who are selected.

Table 3 - Probability of entering the selection (marginal effects)

	All sample		Sample with 8 top 8 pictures	
	Equality	Minority	Equality	Minority
White face	-.014	-.044	.019	-.038
	(.040)	(.041)	(.045)	(.042)
Top 8	.394***	.444***	.447***	.456***
	(.026)	(.028)	(.030)	(.030)
White face & White subject	.025	.003	.030	.000
	(.038)	(.025)	(.046)	(.027)
pseudo R-squared	.12	.16	.16	.16
n. obs	1440	1344	1008	1152
n. subjects	60	56	42	48

Probit estimates - Standard errors clustered by face

4.5 A measure of discrimination

Since we have detailed information on values, we can calculate more precisely the discriminatory implications of cognitive limitations in the treatment *Equality*. That is, even though the faces do not correspond to people actually incurring the consequences of the decisions, we can calculate gains and losses for each racial group of faces. Specifically, we can distinguish between three types of outcomes:

(A) faces with a value *within* the top 8 and *included* in the selection

(B) faces with a value *lower* than the top 8 and *included* in the selection

(C) faces with a value *within* the top 8 and *excluded* from the selection

For each combination of racial group of faces and subjects, I calculate the total value associated with each of these three categories and the total value corresponding to the optimal selection (sum of the values of pictures within the top 8). The ratio between the first sum (corresponding to the actual selection) and the second (corresponding to the optimal selection) provides an indication of the relative gains and losses for each racial group of faces due to the decisions of each racial group of subjects. Then I calculate for each racial group the ratio of the total of the values of faces included in the selection (whether they belong to the top 8 or not) and the total of the values of the faces inclued in the top 8. This gives a measure of the overall losses and gains incurred by each racial group.

The results are presented in Table 1. Not surprisinly, the losses associated with exclusions of top 8 faces are larger than the gains associated with inclusions of faces outside the top 8). As a group, the white faces receive as much as their potential when they are in the choice sets of White subjects (99%), while they suffer a 19% loss when they are in the choice set of East-Asians. The relative losses are more comparable in magnitude for White and East-Asian faces part of the choice sets of East-Asian subjects (14% and 11% respectively). Thus, overall, the cognitive biases identified here lead to an overall *worse treatment* on average of other racial groups, but this effect is particularly pronounced for East-Asian faces in the choice sets of White subjects.

Table 4 - Losses and gains due to cognitive limitations
Share in Potential Realization

	White Subjects (1)	East-Asian Subjects (2)
Included - Top 8 (A)		
White top 8 selected	.75	.62
East-Asian top 8 selected	.57	.65
Included - Outside Top 8 (B)		
White not top 8 selected	.24	.24
East-Asian not top 8 selected	.24	.24
Excluded - Top 8 (C)		
White top 8 not selected	.25	.38
East-Asian top 8 not selected	.43	.35
Total (A)+(B) White faces	.99	.86
Total (A)+(B) East-Asian faces	.81	.89

5 Discussion and conclusion

This study highlights cognitive limitations in the recall of identities and payoff-relevant information and shows that these limitations are more pronounced across races than within race. These limitations lead to unequal treatment across races: negative discrimination at the top of the value distribution and positive discrimination for those in the middlle. This experiment is a first attempt to shed light on possible implications of cognitive limitations in re-identification for social interactions. As discussed in the introduction, these biases may have salient implications for the structure of cross-racial relations. Like a language, a re-identification technology enhances coordination of decisions and efficiency and seem to be developed very early on in childhood. In this context, understanding how re-identification technologies develop and how they shape social relations could deepen our understanding of racial integration.

The current study provides a benchmark, and aims primarily at drawing attention to the phenomenon and its relevance for economic interactions. There are many ways forward from here. A first important extension is to study how these biases vary across ethnic groups and social environments. Second, the experiment conducted here is in a non-strategic setting. One obvious extension is to study the implications of such biases in the context of repeated strategic interactions, for example, in a trust game. One could investigate whether these biases lead to differences in the sustainability of cooperation within and across races. The advantage of an experimental setting would be to be able to attribute faces (and races) to subjects exogenously, and thereby break the link between facial features and payoff-relevant information. It is not possible to make inferences about the trustworthiness of a person based on facial features. This is a great advantage in order to isolate the role of memory from other possible cognitive mechanisms affecting decisions such as perceptions and stereotypes. For example, one could attribute faces of different races to subjects of different races, and study how cooperation develops within and across races. Moreover, in a strategic setting, subjective beliefs about other players' abilities to remember may play a role as well.[13] Why bother being cooperative if the other players will not remember?

Also, the study focuses on faces as identity markers. Faces are obvious reliable signals of identity, a face is generally unique to an individual and it cannot be altered or copied easily (Gambetta and Bacharach (2001)). Faces clearly play a critical role in many social interactions. This said, natural extensions to this study would be to investigate whether similar biases arise with other reidentification technologies, names being the other obvious candidate. Like faces, names typically have a cultural imprint, such that similar biases may arise. One could even investigate whether multiple identifiers (faces and names for example) aggravate the biases or not.

Finally, one could also study how these biases interfere with other cognitive biases, such as stereotypes. Perceptions about probable relevance for future interactions, abilities or trustworthiness may affect how much effort is put into remembering someone. For example, Burgess and Greaves (2009) show that teachers' perceptions about the quality of pupils also suffer from biases along

[13] This point was kindly brought to my attention by Robin Cubitt

the racial dimension. It could be that these biases in perceptions feedback into biases in memory.

To my opinion, these questions deserve attention and careful analyses, in the laboratory and also possibly in the field. The main challenge of field experiments on this phenomenon will be to be able to find reliable measures of payoffs and possibly even measures of perceptions of these payoffs.

These challenges are left for future research.

References

[1] Aigner, D.J. and G.G. Cain (1977), Statistical Theories of Discrimination in Labor Markets, *Industrial and Labor Relations Review* 30 (2), 175-187.

[2] Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1998). Speech and auditory processing during infancy: Constraints on and precursors to language. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* 2, 147–254, New York: Wiley.

[3] Axelrod, R. and W.D. Hamilton (1981), The Evolution of Cooperation, *Science* 211, 1390–96.

[4] Basu, S., J. Dickhaut, G. Hecht, K. Towry and G. Waymire (2009) Record-keeping alters economic history by promoting reciprocity, *Proceedings of the National Academy of Science* 106(4), 1009-14.

[5] Burgess, S. and E. Greaves (2009), Test Scores, Subjective Assessment and Stereotyping of Ethnic Minorities, CMPO Working Paper 09/221

[6] Chiroro J and T.Valentine (1996), An investigation of the contact hypothesis of the own-race bias in face recognition. *Quarterly Journal of Experimental Psychology.*48A, 879–894.

[7] Currarini, S., M.O. Jackson & P. Pin (2009), An Economic Model of Friendship: Homophily, Minorities, and Segregation, *Econometrica*, 77(4), 1003-1045.

[8] Cunningham W.A., M.K. Johnson, C.L. Raye. J.C. Gatenby, J.C. Gore and M.R. Banaji (2004), Separable neural components in the processing of Black and White faces. *Psychological Science.* 15, 806–813.

[9] Duarte, J., S. Siegel and L.A. Young (2009), Trust and Credit. Available at SSRN: http://ssrn.com/abstract=1343275

[10] Duchaine,B., Yovel,G. (2008). Face Recognition. in Basbaum,A., R. Hoy, A. Kaneko, G. Shepherd and G. Westheimer (ed.) *The Senses: A Comprehensive Reference*. Amsterdam: Elsevier, 1st edition, 329-358. ISBN: 0-12-639482-2.

[11] Eckel, C. and R. Petrie (2010), Face Value, *American Economic Review* forthcoming.

[12] Ferguson D.P., G. Rhodes and K. Lee (2001) They all look alike to me: Prejudice and cross-race face recognition. *British Journal of Psychology*;92:567–577.

[13] Fong E and W.W. Isajiw (2000), Determinants of friendship choices in multiethnic society, *Sociological Forum* 15, 249–269.

[14] Fryer, R. and M.O. Jackson (2008), A Categorical Model of Cognition and Biased Decision Making, *The B.E. Journal of Theoretical Economics* (Contributions) 8 (1), Article 6.

[15] Gambetta, D. and M. Bacharach (2001), Trust in signs, in K. Cook (ed.) *Trust and Society*, New York: Russell Sage Foundation, 148-184.

[16] Golby A.J., J.D. Gabrieli, J.Y. Chiao and J.L. Eberhardt (2001). Differential responses in the fusiform region to same-race and other-race faces. Nature Neuroscience 4, 845–850.

[17] Hanley, J.R. and G. Cohen (2008), Memory for People: Faces, Names and Voices in Cohen, Gillian and Conway, Martin A., *Memory in the Real World*, Psychology Press, UK

[18] Ito T.A., G.R. Urland (2003), Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology* 4, 616–626.

[19] Kuhl, P.K. (1993), Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics* 21, 125–139.

[20] Kurzban, R., J. Tooby and L. Cosmides (2001). Can race be erased?: Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, 98(26), 15387-15392

[21] Lavrakas P.J., J.R. Buri and M.S. Mayzner (1976), A perspective on the recognition of other race faces. *Perception & Psychophysics* 20, 475–481.

[22] Levin D.T. (1996), Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22, 1364–1382.

[23] Levin D.T.. (2000) Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology*: General 129, 559–574.

[24] Maclin O.H. and R.S. Malpass (2003), The ambiguous-race face illusion. *Perception* 32, 249–252.

[25] McPherson M, L. Smith-Lovin and J.M. Cook (2001). Birds of a feather: homophily in social networks. *Annual Review of Sociology* 27, 415–444.

[26] Meissner C and J. Brigham (2001), Thirty years of investigating the own-race bias in memory for faces. *Psychology, Public Policy, and Law*.7, 3–35.

[27] Mollica, k.A., B. Gray and L.K. Treviño (2003), Racial Homophily and Its Persistence in Newcomers' Social Networks, *Organization Science*14 (2), 123-136.

[28] Montepare JM, Opeyo A (2002). The relative salience of physiognomic cues in differentiating faces: A methodological tool. *Journal of Nonverbal Behavior*. 26, 43–59.

[29] Mullhainathan, S. (2002), A Memory-Based Model Of Bounded Rationality, *The Quarterly Journal of Economics* 117(3), 735-774.

[30] Mullainathan, S., J. Schwartzstein and A. Shleifer (2008), Coarse Thinking and Persuasion, *The Quarterly Journal of Economics* 123(2), 577-619, 05.

[31] Ng W, Lindsay R. (1994) Cross-face recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology* 25, 217–232.

[32] Pascalis, O., M de Haan, C.A. Nelson (2002), Is Face Processing Species-Specific During the First Year of Life? *Science*.296 (5571), 1321 - 1323

[33] Rhodes G, S. Brake, S. Tan and K. Taylor (1989) Expertise and configural coding in face recognition. *British Journal of Psychology* 80, 313–331.

[34] Rule, N. O. and N. Ambady (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science*, 19, 109-111.

[35] Shepherd, J.W., F. Gibling and H.D. Ellis (1991), The Effects of Distinctiveness, presentation time and delay on face recognition, *European Journal of Cognitive Psychology 3,* 137-45.

[36] Slone A, J. Brigham and C. Meissner (2000), Social and cognitive factors affecting the own-race bias in Whites. *Basic and Applied Social Psychology*;22:71–84.

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

[38] Taylor, S., S. Fiske, N. Etcoff, and A. Ruderman (1978) Categorization and Contextual Bases of Person Memory and Stereotyping. *Journal of Personality and Social Psychology* 36, 778-793.

[39] Tibbetts, E. A. and J. Dale (2007), Individual Recognition: It is Good to be Different, *Trends in Ecology and Evolution* 22 (10), 529-37

[40] Todorov, A., A.N. Mandisodza, A. Goren and C.C. Hall (2005), Inferences of competence from faces predict election outcomes. *Science*, 308, 1623–1626.

[41] Todorov, A. and B. Duchaine. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology* 25, 395-410.

[42] Valentine, T. (1992), Towards an exemplar model of face processing The effects of race and distinctiveness, *The Quarterly journal of experimental psychology. A, Human experimental psychology* 44 (4) 671-703.

[43] Werker, J.F. and R.C. Tees (1984), Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75, 1866–1878