Cognitive Discrimination - A Benchmark Experimental Study

Michèle Belot*
Nuffield College, University of Oxford

March 16, 2011

Abstract

This study asks the following questions: Is it harder to remember people from other races? And do these cognitive limitations have discriminatory implications? We conduct an experiment in a controlled laboratory environment. Participants are shown pictures of potential "candidates" of different races - East Asian and white - and each candidate is associated with a value. There are no systematic differences in the value distribution across races. They are asked to recall the faces of the candidates with the highest values. I find that people are much better able to recall candidates with higher values if they are of the same race. This leads to positive and negative discrimination at the same time: those at the bottom of the value distribution benefit while those at the top lose out. These results suggest that cognitive biases could play a role in the nature of cross-racial relations, in particular for racial discrimination, homophily and phenomena relying on repeated interactions and individual recognition, such as the formation and maintenance of social ties and the establishment of trust relationships.

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

JEL codes: J71, C91, D83

1 Introduction

Remembering people plays a key role in many social and economic contexts. Most of our daily interactions involve people we have met before and whom we will meet again: colleagues, friends, professional relations, acquaintances, local merchants, etc. Keeping track of people - who they are and what they do - improves efficiency. We do not need to "start over again" every time we meet the same person. In an increasingly racially diverse environment, it is relevant to ask: are we better at keeping track of people of our own race? For example, if asked for a recommendation for a job, are we less likely to remember suitable people from other racial groups?

Of course there are many reasons why we may remember better some people than others: the frequency of interactions, our perceptions about them, etc. The question is whether on top of that biases in the *ability* to memorise and retrieve information about others may play a role as well. For economists the relevant question is how do people record and retrieve information of the type "Person x has productivity y". Identification involves mapping an identifier - a name, a face, etc. - to economically relevant information. For example, a job recommendation often involves remembering a name and a person's qualifications. If there are biases in memory, this could have important implications in many dimensions of life, for example for whom we form social ties with, whom we trust or whom we recommend for a job. There are many economic transactions based on repeated interactions where individual recognition matters. Limitations in memory could have discriminatory implications even in the absence of any relevant differences across races - for example in productivity distributions - and in the absence of any race-specific preferences or stereotypes.

This study provides a first set of experimental results regarding the joint recall of identities and economically relevant information, and investigates whether biases arise along the racial dimension. It is designed to serve as a benchmark and is conducted in a simple and neutral environment. The focus of this study is on facial individual recognition. Experimental participants see people, learn information about them in a controlled manner and then make a decision requiring recall. We provide a clean measure of the efficiency and discriminatory implications of memory limitations.

Facial recognition plays a key role in many social interactions. Often a face is all we rely upon to identify someone. Local merchants typically know their customers only by face, not by name or any other identifier. Professional conferences, business social events, interactions between teachers and students, or between students themselves, are all examples of environments with repeated encounters and where facial recognition is an essential technology used to identify others. But obviously, there are other technologies used to identify people - names in particular. What is important to note is that these technologies are likely to share common features, insofar as they are likely to be shaped by the environment we grow up in and by the nature of our social interactions. The general question to ask is whether the technologies we use to rembember and record information about others are more effective for people of our own race or culture than for other people and, if that is the case, how does this affect cross-racial and cross-cultural relations?

The idea that memory limitations in individual recognition may affect economic transactions has been almost unexplored in economics. This is surprising given the large interest in phenomena such as racial discrimination and homophily (the tendency to associate with people of the same race). To

3

explain these phenomena, the literature in economics has mainly focused on the role of preferences (taste-based discrimination) as proposed by Becker (1961) and beliefs (statistical discrimination) as proposed by Arrow (1973). The implications of a cognitive racial bias in person recognition are related to those present in contexts of statistical discrimination. In these models, discrimination may arise in the absence of differences in the underlying distributions of characteristics (e.g. productivity) - through the asymmetry in information. Aigner and Cain (1977) and Cornell and Welsh (1995) propose models where discrimination arises through biases in the technology used to *screen* people and evaluate their productivity. They conjecture that employers may be better at *evaluating* the productivity of workers from their own race than other races and this implies that the information is noisier for people of other races. The result is a compression in the distribution of rewards: there is positive discrimination at the bottom of the productivity distribution and negative discrimination at the top. Similarly here, memory limitations result in noisier information. It can even be the case that people are perfectly informed about relevant characteristics, but the key aspect is remembering these characteristics and attributing them to the right person. The most relevant work in that regard is by Fryer and Jackson (2008), who propose a model of discrimination precisely based on bounded memory.[1] In their model people are sorted into categories and each category has a prototype - a unique vector of attributes. People keep track of the variation in attributes *across* categories but not within a category. People sorted into the same category are blended together. For example, they could be sorted into two categories: high-productivity and low-productivity people.

---

[1] This work fits in the literature on bounded memory, and coarse thinking see Mullainathan (2002) and Mullainathan et al. (2008).

They argue that minority groups may be sorted into coarser categories than majority members because they are less likely to be involved in frequently repeated interactions. Again, this leads to positive discrimination for those at the bottom and negative discrimination for those at the top.

The question raised here is how people record and retrieve information involving an association between "identifiers" on the one hand (attributes that enable identification) and payoff-relevant information on the other. As we will discuss in the next section, we know that own-race biases arise in the ability to memorise facial attributes. But the question of whether biases and inefficiencies arise in information recall critically depends on the precision by which payoff-relevant information is recorded. Only if it is recorded sufficiently precisely, biases will arise as well.

The experimental study proceeds as follows. Participants are presented with a sequence of 24 pictures of East Asian and white people - whom I call "candidates." Each candidate is associated with a value. In the second stage they are asked to select 8 people among the candidates they have seen. They see pictures again but they do not see the value. Their earnings depend directly on the values of the candidates they select. These values have a neutral frame. They are a general measure of payoff-relevance and could be interpreted in different ways. In repeated games the value of a person represents the benefit from the interaction (and could summarise the relevant past actions of the player). In network formation games they represent the value from linking with a person and in labour markets they represent a measure of productivity. Moreover, these values are exogenously assigned by a random technology and there is no difference in the distribution of values across racial groups.

A controlled environment presents two major advantages 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 candidate on the picture) does not incur any loss or 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 payoffs associated with each candidate are fully controlled for and participants are fully informed about these individual payoffs and about the procedure of assignment of payoffs to candidates. This exogenous assignment means that the faces of candidates do not contain any payoff-relevant information. This is important given the recent evidence showing relationships between facial features (and in particular race) and inferences about personal characteristics such as competence and trustworthiness (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 that can correctly map faces to values is memory.

The baseline treatment presents participants with an equal number of candidates of each race. The data show a clear own-race bias in recall. This bias is most pronounced among white participants, who recall high-value candidates from their own race much more accurately than they recall high-value candidates from other races (66% of the white candidates with a value among the eight highest values are selected against 50% for East-Asian candidates). Second, participants select an equal number of candidates of each race - 4 on average. Thus, we find evidence of positive and negative racial discrimination at the same time. Higher ranked candidates of the other race are less likely to be selected, but the opposite is true for lower ranked candidates.

6

I also investigate an interesting dual aspect of cross-race re-identification, which is that race might enhance identification when there are few people of a particular race (in a crowd of white people, an East-Asian man will stand out). Race has been found to be a prime characteristic encoded about others (Montepare and Opeyo (2002)). It is a *distinctive* attribute - and distinctiveness has been found to enhance re-identification (Shepherd et al. (1991), Tibbetts and Dale (2007)). To shed light on this dual aspect of cross-racial re-identification, I implement a second treatment where there are only two East-Asian candidates, a man and a woman. In this case I find that decisions involving East-Asian candidates are significantly more efficient than those involving an equal number of candidates of each race, and this is true both for white and East-Asian participants.

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

2   The Own-Race Bias in Facial Recognition

There is a well-established literature in psychology on the "own-race bias" in facial recognition - the fact that people are better at remembering faces of people of their own race than other races (see Meissner & Brigham, 2001; Slone, Brigham, & Meissner, 2000 for reviews). Psychologists have repeatedly stressed the implications of such bias for criminal convictions based on eyewitness identification. However, as we mentioned earlier, in a social or economic context, the critical question is not whether one recalls a face or a name. The relevant question is whether one recalls *who a person is*

- that is, whether one can retrieve relevant *information* associated with a face or a name - and whether there are biases in the accuracy of recall of such information. Hanley (2008) documents the many situations in which people report being able to remember faces but cannot recall the context. We do not know how people memorise payoff-relevant information and map this information to the corresponding identifiers. As we shall argue later, it is not clear at all that own-race biases should be as large or even arise. But obviously, the findings on the own-race bias in facial recognition are relevant for our study, so we review this literature in some detail.

Psychologists and neuropsychologists have extensively studied 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 participants. The overwhelming consensus among social psychologists is that an own-race bias exists and is 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 specific 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 obvious 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

9

processing of faces of different races follows a similar developmental course.

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 - how people record and retrieve information of the type "Person x has productivity y". The most relevant 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 study how participants recall the contents of interactions between people of mixed gender and race ("*Who said What?*")[2]. They show that participants are more likely to misattribute statements of people of the same race than different races. To the best of our knowledge, there are no studies providing evidence on how people memorize identities and payoff-relevant information. Payoff-relevance introduces a specific nature to information: it is a cardinal measure, which possibly commands a specific way of recording and organising information. As we shall argue later, it is not clear that own-race biases in information recall would necessarily arise.

3   Experimental Design

The experiment was conducted at the laboratory of the Nuffield Centre for Experimental Social Sciences (Oxford) in September and November of 2009 and March 2010. 116 participants were recruited: 61 Caucasians and 55 East Asians. Invitations were sent by e-mail to participants in the pool with East Asian and British last names, without mentioning race or ethnicity[3]. Sessions were also relatively small in size (maximum 15 participants at a time)

---

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

[3]Participants were recruited using ORSEE (Greiner, 2004). 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).

and the split across race was not equal for each session. The experiment lasted for 30 minutes in total, including a post-experimental questionnaire asking information about ethnicity, age, occupation, country of birth, age of arrival in the UK and a self-assessment of ability to remember faces of people in general. Participants received a £4 show-up fee and an additional payment depending on the performance in the memory task (see explanation below), bringing the total payment to £8.60 on average.

The experiment is structured in 3 stages. The first stage is a *viewing stage* where participants view an automated sequence of 24 pictures. Each picture is shown for 3 seconds and then the screen moves to the next picture.[4] Each picture appears together with its value. The second stage is a *selection stage* – participants see the same 24 candidates again, but without their associated values, and are asked to select 8 candidates that will enter a lottery (in a third stage). The lottery determines their final earnings. Importantly, pictures appear in a different sequence, and from a different angle. The sequence is not automated at this stage and participants can go back and forth between pictures for 3 minutes. In case of incomplete selection (less than 8 candidates selected by the subject), candidates are selected at random among the remaining non-selected candidates to complete the selection (it turns out that in the experiment all participants without exception selected all 8 pictures themselves). Note that the sequence of presentation of candidates is randomized for each participant, and for each stage. The third stage is a *lottery*, whereby one of the candidates from the selection in stage 2 is picked at random.[5] The subject's earnings are equal to the value

---

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

[5] 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,

of the picture divided by 10, in British pounds.

Pictures of candidates

Pictures of candidates were drawn from a database provided by TAR-RLAB [6]   These pictures show only the face of the person. Pictures were selected according to a number of criteria to guarantee homogeneity in shooting conditions[7]. The database contains 11 East Asian men, 15 East Asian women, 35 Caucasian men and 44 Caucasian women.

For each subject, a set of 24 candidates was randomly chosen. Two pictures of each candidate from a different angle were randomly chosen (one used in the viewing stage and the other used in the selection stage). This ensures that the task involves face recognition rather than picture recognition and prevents participants 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 candidates

A unique value was pre-assigned randomly to each candidate. The values correspond to random draws from a discrete normal distribution truncated at 10 and 70, with mean 40 and standard deviation 15. This range has been

---

there were always 8 pictures entering the lottery. In practice, all participants selected 8 pictures exactly.

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

[7]All images were extracted from standard digital video (720x480), with the background removed and the faces scaled to be roughly equated in terms of size. The pictures were selected according to the following parameters: Race (East-Asian 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: 0°, 15°left, 15°right, 30°left, 30°right.

chosen so that all values have 2 digit numbers. The choice of the normal distribution is motivated by the relatively low presence of extreme values in many real applications - extremes that are often those one would like to remember in taking decisions.

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

Treatments

Two treatments are conducted:

Treatment 1 - "Equality treatment" - Choice sets are composed of an equal number of East-Asian and Caucasian candidates: 6 men and 6 women of each race.

Treatment 2 - "Majority-Minority treatment" Choice sets are composed of 22 Caucasian candidates (11 men and 11 women) and 2 East-Asian candidates (1 man and 1 woman).

4   Analysis

4.1   Predictions

It is useful to outline a number of implications of different models that could drive selection decisions in this context. I describe the implications associated with a memory-based model in the spirit of Fryer and Jackson (2008)

and I contrast these implications with those from models based on preferences or stereotypes. I am specifically interested in implications regarding the probability of entering the selection and how this probability varies with the value attached to the picture.

Bounded memory

The simplest way to describe the decision problem is as follows. Participants see a sequence of candidates with values x_i, $i = 1, ...24$. They know the distribution of values attached to these pictures (mean 40 and standard deviation 15). We depart from Fryer and Jackson insofar as we need to distinguish between observable attributes on the one hand - attributes that will be observed at the decision stage - and payoff-relevant information that will not be observed at decision stage. Let us assume that the decision maker forms categories of people based on observable facial attributes: for example, according to race, gender, colour of the hair or eyes. Suppose this results in K equal size categories for the own race and J equal size categories for the other race, with $J < K$. The decision maker cannot distinguish between people sorted into the same category.

In addition to keeping track of the observable attributes uniquely identifying each category, the decision maker also needs to keep track of the payoff-relevant information. As a first step, suppose that she keeps track of the mean value corresponding to each category. In a small sample, categories will have different expected values. Note that one interesting empirical implication is the fact that candidates with very low values could have a positive probablity to be selected as well, because their observable attributes could sort them into a category with a high expected value. We find that this is indeed occurring (Fig. 2).

14

In this framework, the optimal selection consists in first choosing candidates sorted into the category corresponding to the highest expected value. Since categories for the own-race candidates are finer, it is more likely that the category with the highest expected value will be of the own-race. If fewer than 8 candidates belong to this category, all candidates of this top category should be included. The second step consists of comparing the expected value of the second highest category of the own race with the highest category of the other race and select candidates from the category with the highest expected value. The procedure is repeated up to the point where 8 candidates have been selected.

A first insight from this model is that own-race candidates should not necessarily be *more* likely to enter the selection. It could actually be optimal to select more candidates of the other race. To understand why this is the case, consider the following example. Suppose that for each white candidate, there is a corresponding East-Asian candidate with the same value. Suppose further that people form 4 categories of 3 candidates for their own race (with for example expected values of 20, 30, 35 and 55 and 2 for the other race with 6 candidates in each category (for example categories with expected values of 45 and 33 respectively). The highest category of own race has the highest expected value, so all 3 candidates from that category should be selected. Then the second category of own-race candidates has a lower expected value than the expected value of the top category of the other race 45. Thus, it is optimal to select 5 candidates at random among the top category of the other race. One could construct alternative examples, where the optimal composition holds more or fewer candidates of each race. The important point is that bounded memory does not necessarily imply that people should select a lower number of candidates of the other race.

15

A second insight from this model is regarding the relationship between the probability of entering the selection and the value. Since the probability of entering the selection should be identical for all candidates sorted into the same category, the coarser the categories the flatter the relationship between values and probability of being selected. If there is no memory and just one category, then the probability of selection is identical for all candidates. In the case of perfect memory, only the candidates belonging to the top 8 should be selected while those outside the top 8 should never be selected. If we only have two categories, then there is a positive probability that a candidate with a value belonging to the 8 eighest values is sorted into the category with lower expected value. This is why the relationship becomes flatter.

The predictions so far rely on a specific structure of bounded memory, in particular, we have assumed that people compute and record correctly the mean value associated with the category. Since the exercise here involves a mapping of identifiers to information, it seems reasonable to assume that the information may also be recorded in a coarse manner. Both recording the individual information and aggregating it to the category could involve approximations and coarsening of information. For example, people could use simple heuristics such as "above the mean", or "above 55" and associate these labels to categories of identifiers. It is obvious that coarser recording of payoff-relevant information will wash away part of the biases predicted above. Take for example the extreme case where the information does not get recorded at all, then of course better facial recognition provides no advantage in better identification of those with a higher payoff. Own-race biases would also not survive here to heuristics such as "above the mean".

16

Biases will arise if the information is recorded in a sufficiently precise manner. But given the difficulties people report having in remembering contexts and information associated with faces, it is a priori not clear that own-race biases in information recall will be strong or even arise.

.Preferences or Stereotypes

A second possible model is one that would allow a role for preferences or stereotypes. Becker's (1961) model of taste-based discrimination conjectures that discrimination arises through preferences for interacting with people of one's own racial group. Preferences and tastes should have a limited role here by design since there is no "interaction" between the decision-maker and the person who is involved in the decision (the candidate on the picture). There is no other benefit from selecting a particular picture other than than the economic value attached to that picture. Also, decisions only affect the decision-maker and not the candidates, such that other-regarding considerations should not affect decisions. Yet participants come to the lab with a real world experience. Preferences may not be totally turned off and may affect decisions even in this environment. Preferences for people of similar race are equivalent to attributing a lower value to people of the other race. Consequently, the probability of selection, conditional on value should be lower for other-race candidates than own-race candidates. That is, across the board, we should see that all candidates of the other race are less likely to enter the selection.

The second model of discrimination is a model of rational or statistical discrimination based on negative stereotypes (Phelps, 1972 and Arrow, 1973). Stereotypes can affect decisions in environments where there is uncertainty about the value of a person. Here the environment rules out this

17

Table 1: Treatments and number of participants

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

uncertainty by design: Participants are perfectly informed about the values, and about the distributions of these values. Yet, again, participants may bring stereotypes formed outside the lab into the lab. The implications of negative stereotypes are identical to the implications of same race preferences. They also come down to attributing a lower value to people of the other race and should also lead to a lower probability of selection conditional on value. This implication contrasts with the implications of a memory-based model.

4.2 Analysis and results

4.2.1 Summary statistics

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

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

Table 2 - participants summary statistics

	Caucasian White participants		East-Asian participants	
	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)

Overall, participants did significantly better than chance, with a proportion of correctly allocated candidates equal to 61% ($P = .000$)[8]. The mean value of the selected candidates is 48 (S.E. .47), significantly higher than the benchmark for chance (40, $P = .000$).

4.2.2 Treatment *Equality*

We start with a simple figure (Fig. 1) presenting the average proportion of top 8 candidates selected by racial groups of participants and candidates. The selection shows a substantial own-race bias[9]. The proportion of top 8 East-Asian candidates selected by white participants is .50 against .66 of top 8 white candidates ($P = .001$). The opposite trend is observed for East-Asian participants: a proportion .65 of top 8 East-Asian candidates is selected against .56 of top 8 white candidates), but the difference is not statistically significant ($P = .25$).

I now turn to a more detailed analysis of the selection decision. I estimate a probit model for the selection decision with the objective of describing

---

[8]Two-sided Mann-Whitney tests with the individual proportions as independent observations.

[9]All reported tests correspond to post-estimation F-tests based on linear regression estimates pooling all choices made by the 116 participants and clustering the standard errors by faces.
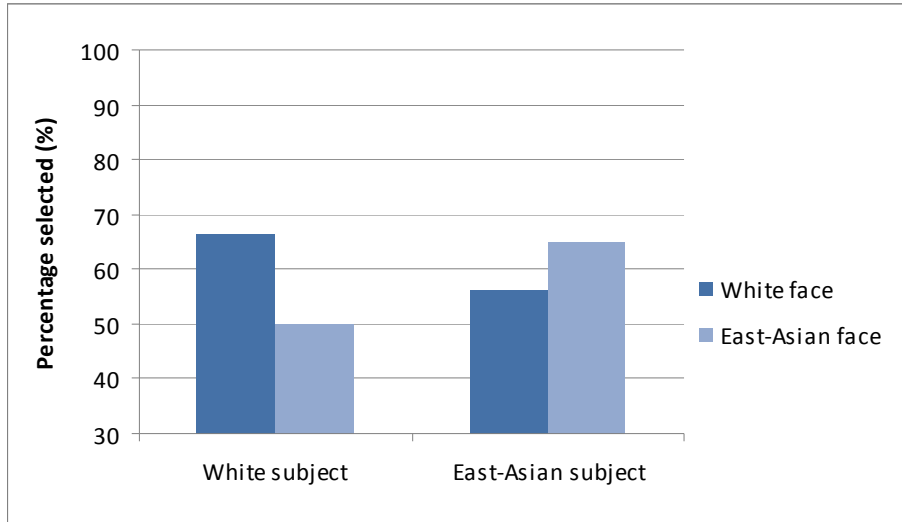
Figure 1: Percentage of pictures selected among the top 8 - Treatment *Equality*

more precisely how and where these biases arise. One immediate question is whether the inefficiencies in allocation decisions arise at the top of the distribution of the top 8 candidates - 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, I distinguish between the first four values and the values ranked 5th to 8th. The results are shown in Table 3. There are three conclusions to draw from these results. First, among the set of candidates 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. Second, this effect does not eliminate cross-racial biases. Higher ranked candidates (belonging to the top 4) of the other race are significantly less likely to enter the selection. There is a difference between racial groups of participants regarding the selection of the lower ranked can-

20

didates (5th to 8th) of the other race. While for white participants, there is no difference between high ranked and lower ranked values, there is a difference for East-Asian participants. The bias seems especially present for higher ranked candidates (although the difference in probabilities of selection between lower ranked and higher ranked candidates is not significant). Third, conditional on being in the top 8 or not, there is no difference in the probability of entering the selection according to race. That is, both white and East-Asian participants choose an equal number of candidates of East-Asian candidates and white candidates. This is important because it implies positive and negative discrimination at the same time: high-ranked candidates from the other race are less likely to enter the selection, but this benefits lower-ranked candidates (that do not belong to the top 8). Coming back to the model of bounded memory, an equal number of pictures of each race can be optimal but then if there are indeed coarser categories for the other race, the average value of the selected candidates of the other race should be lower than the average value of the selected own-race candidates. This is exactly what happens. This result contrasts with what we would expect if negative stereotypes of preferences for own-race candidates would drive decisions. Both mechanisms have similar implications and should lead to, all else equal, a lower probability of entering the selection. We find that this is not the case. On the contrary, this result is perfectly consistent with the hypothesis of bounded memory.

To get more insight in the selection process, Fig. 2 and 3 plot the frequency of selection as a function of values. The prediction is that coarser memory for the other race should lead to a flatter relationship between the values and the probability of selection. These figures show that the slope of

Table 3 - Probability of entering the selection - Treatment "Equality"

	White part.	EA part.	White part.	EA part.
other race	-.065 (.044)	.008 (.052)	.065 (.044)	.008 (.053)
top 8	.486 (.048)***	.458 (.053)***		
other race × top 8 (i)	-.192 (.054)***	-.085 (.069)		
rank 1 to 4 (i)			.567 (.051)***	.608 (.054)***
other race × rank 1 to 4 (ii)			-.179 (.066)**	-.159 (.075)*
rank 5 to 8 (iii)			.433 (.065)***	.346 (.074)***
other race × rank 5 to 8 (iv)			-.191 (.059)***	-.030 (.095)
(i) = (iii) p-value			.067	.000
(ii) = (iv) p-value			.889	.210
n.obs	768	672	768	672
n. participants	32	28	32	28
Pseudo R^2	.13	.14	.13	.15

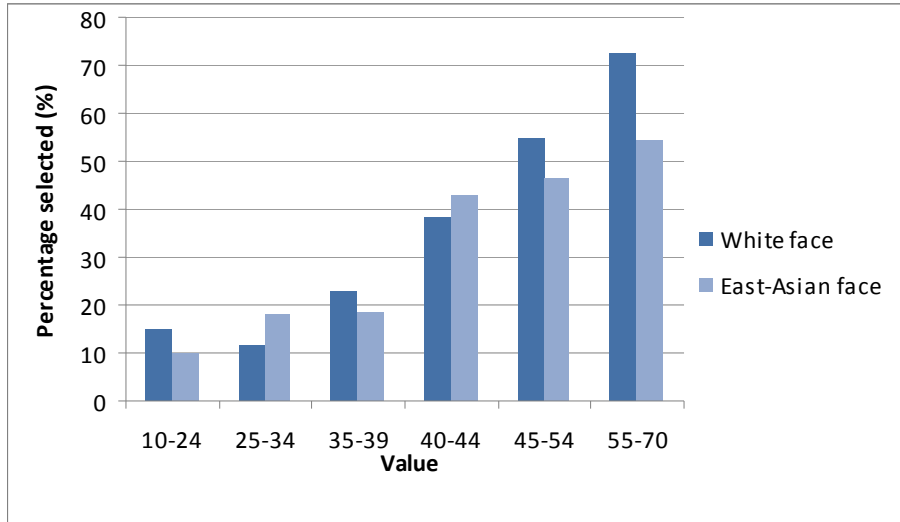Probit estimates - Marginal effects - Standard errors clustered by candidate

Figure 2: Values and frequency of selection - White subjects

the relationship between the value and the probability of selection is steeper for own race candidates than other race candidates. We test this hypothesis more formally in Table 4 where we interact dummies corresponding to intervals of values with candidate race. We find that indeed values decrease the probability of selection but less so for candidates of the other race (all interaction dummies are positive). We conclude that the results are consistent with coarser memory for the other race.

Summarizing the results, the evidence shows an own race bias when there is an equal number of candidates of each race. This effect is more pronounced among white participants than East-Asian participants.

In that respect, it is worth pointing out though that East-Asian participants are living in the UK, and therefore might be a selected sample with respect to the ability to distinguish white candidates.

Table 4 - Probability of entering the selection - Treatment "Equality"

	White subject	East Asian subject
other race	-.137 (.080)*	-.121 (.095)
value 55-70		
value 40-54	-.199 (.066)***	-.339 (.064)***
value 25-39	-.463 (.049)***	-.507 (.051)***
value 10-24	-.354 (.034)	-.39.21 (.043)***
value 40-54 × other race	.153 (.111)	.068 (.110)
value 25-39 × other race	.201 (.115)*	.224 (.113)**
value 10-24 × other race	.056 (.149)	.028 (.145)
n.obs	768	672
n. participants	32	28
Pseudo R^2	.14	.14

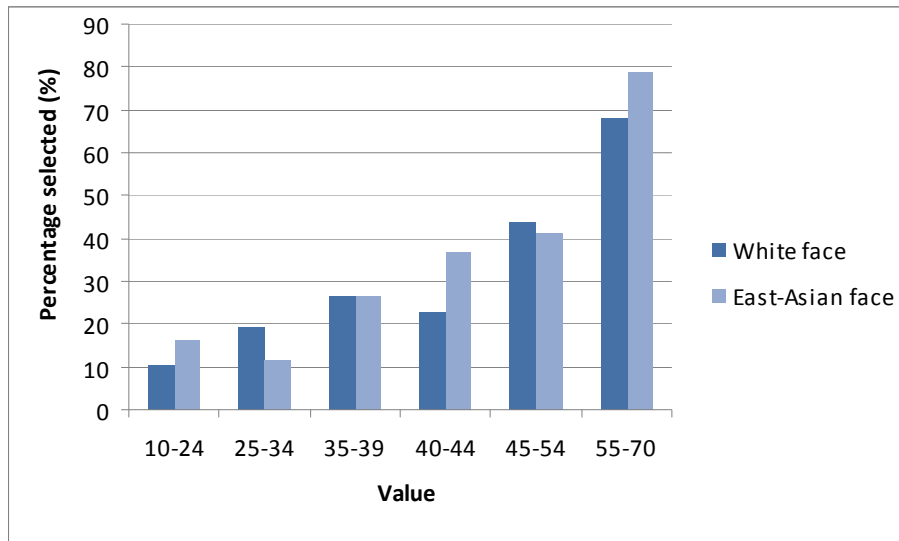Probit estimates - Marginal effects - Standard errors clustered by candidate

Figure 3: Values and frequency of selection - East-Asian subjects

4.2.3 Treatment *Minority*

In the second treatment, the composition of facial stimuli only includes 2 East-Asian candidates, 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). The bias now takes the form of a *minorty bias*: candidates from the minority group are more likely to be correctly allocated than candidates from the majority group. Table 5 reports probit estimates testing for treatment effects for white and East Asian participants separately. The results show a substantial and significant improvement in the probability of entering the selection for top 8 East Asian candidates. The treatment effect is comparable in magnitude for both white and East-Asian participants. Also, notably, there remains no
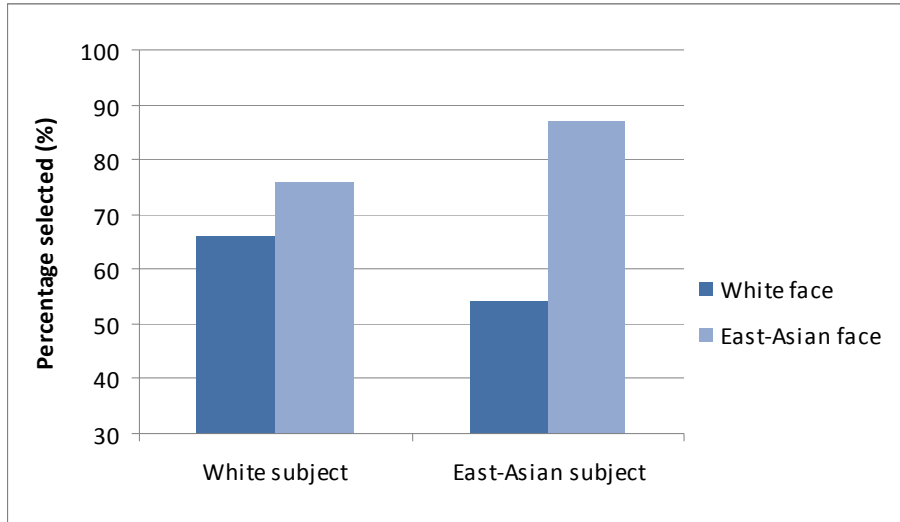
Figure 4: Percentage of pictures selected among the top 8 - Treatment *Minority*

systematic difference in the overall conditional probability of entering the selection across racial groups of candidates. That is, on average, white and East-Asian participants select the same number of white candidates (7.33) and the same number of East-Asian candidates (0.66).

4.2.4   A measure of discrimination

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

Table 5 - Probability of entering the selection - Treatment "Minority"

	White subj.	EA subj.
Asian candidate	-.046 (.043)	.017 (.050)
Top 8	.485 (.048)***	.373 (.051)***
Asian candidate × Treatment minority	-.017 (.032)	.027 (.046)
Asian candidate × top 8	-.171 (.052)***	.103 (.082)
Top 8 × Treatment minority	.016 (.069)	-.042 (.062)
Asian candidate × top 8 × Treatment minority	.255 (.127)**	.334 (.149)**
n.obs	1464	1320
n. participants	61	55
Pseudo R^2	.16	.13

Probit estimates - Matginal effects - Standard errors clustered by candidate

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

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

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

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 associated with candidates within the top 8) for each racial group of participants and candidates. 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 candidates due to the decisions of each racial group of participants. Then I calculate for each racial group the ratio of the sum of values of candidates included in the selection (whether they belong to the top 8 or not) and the sum of values associated with the candidates included 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 6. Not surprisingly, the losses associated with exclusions of top 8 candidates are larger than the gains associated with inclusions of candidates outside the top 8. As a group, the white candidates realize 94% of their potential when they are in the choice sets of white participants, while they suffer a 23% loss when they are in the choice sets of East-Asians. The relative losses are comparable in magnitude for white and East-Asian candidates that are in the choice sets of East-Asian participants (20% and 8% 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 candidates in the choice sets of white participants.

Table 6 - Losses and gains due to cognitive limitations
Share in Potential Realization

	White participants (1)	East-Asian participants (2)	Z-test (1)=(2) P-value
Included - Top 8 (A)			
white top 8 selected	.69	.56	.07
East-Asian top 8 selected	.52	.65	.03
Included - Outside Top 8 (B)			
white not top 8 selected	.25	.21	.55
East-Asian not top 8 selected	.28	.27	.91
Excluded - Top 8 (C)			
white top 8 not selected	.31	.44	.07
East-Asian top 8 not selected	.48	.35	.03
Total (A)+(B) white candidates	.94	.77	
Total (A)+(B) East-Asian candidates	.80	.92	

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 lower in the distribution. This experiment is a first attempt to shed light on possible implications of cognitive limitations in 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, technologies used for individual recognition and identification enhance efficiency and seem to be developed very early on. In this context, understanding how these technologies de-

velop and how they shape social relations could deepen our understanding of racial integration.

The current study provides a benchmark and aims at drawing attention to the phenomenon and its relevance for economic interactions. 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 homophily* (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 *preferences* (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.[10] Limitations in the ability to remember people and the resulting uncertainty in the value of continuing a relationship have not been considered at all in this literature.

The second class of applications regards reputation-based mechanisms, relying on the repeated character of interactions and memory, such as trust and reciprocal altruism (Axelrod, 1981). These mechanisms are often very demanding in the information-structure required to sustain cooperation - in terms of players' identities and records of past behaviour. Basu et al. (2009) show that facilitating recordkeeping with an external support (e.g. allowing

---

[10]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.

participants to take notes) enhances cooperation. Cognitive limitations in cross-race re-identification may therefore play a role in hampering cooperation across races and possibly compromising the use of these mechanisms to enforce cooperative behavior across races.

A third class of applications is the use 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. 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. I propose here a novel angle and argue that cognitive mechanisms may play a role in these phenomena and explain possible racial biases associated with them.

There are many ways forward from here. An important first 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. Moreover, in a strategic setting, subjective beliefs about other players' abilities to remember may play a role as well.[11] Why bother being cooperative if the other players will not remember?

---

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

Also, the study focuses on faces as identity markers. Faces are an obvious starting point because they play a key role in many social interactions. A face is generally unique to an individual and cannot be altered or copied easily (Gambetta and Bacharach (2001)). But we use other technologies to identify people and natural extensions to this study would be to investigate whether similar biases arise with other identification 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, the environment I consider is extremely simple. The payoff-relevant information is captured by a single value. It would be interesting to study recall in a more complex environment, where the relevant information takes a more complex structure. Also, the bridge between the experimental task and real world applications is long. One could ask whether these biases subsist in settings where people interact with each other for a longer period and may also have access to external support to memory.

In my opinion, these questions deserve attention and careful analysis, in the laboratory and also possibly in the field. The main challenge of field experiments on this phenomenon is 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] Arrow, K. (1973), The Theory of discrimination, in *Discrimination in Labor Markets*, O. Ashenfelter and A. Rees, eds. Princeton, NJ: Princeton University Press.

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

[6] Becker, G. S. (1961), The economics of discrimination, 2nd Ed. Chicago: University of Chicago Press.

[7] 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.

[8] Cornell, B. and I. Welch (1996), Culture, information, and screening discrimination, *The Journal of Political Economy* 104(3) 542-571

[9] Currarini, S., M.O. Jackson & P. Pin (2009), An economic model of friendship: Homophily, minorities, and segregation, *Econometrica*, 77(4), 1003-1045.

[10] 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.

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

[12] 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.

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

[14] 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.

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

[16] 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.

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

[18] 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.

[19] Greiner, B. (2004), The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics. University of Cologne, *Working Paper Series in Economics* 10.

[20] 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

[21] 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.

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

[23] 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

[24] 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.

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

[26] 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.

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

[28] 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.

[29] 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.

[30] 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.

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

[32] Mullainathan, S. (2002), A memory-based model of bounded rationality, *The Quarterly Journal of Economics* 117(3), 735-774.

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

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

[35] 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

[36] Phelps, E. (1972), The statistical theory of racism and sexism, *The American Economic Review* 62(4), 659-661

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

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] Tibbetts, E. A. and J. Dale (2007), Individual recognition: It is good to be gifferent, *Trends in Ecology and Evolution* 22 (10), 529-37

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

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

[46] 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.

[47] 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