Journal of Behavioral and Experimental Economics 104 (2023) 102008

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Economics

journal homepage: www.elsevier.com/locate/jbee

Remember me? The role of gender and racial attributes in memory☆

Michèle Belot [a], Marina Schröder [b,*]

[a] Department of Economics, Cornell University, 109 Tower Road, 404 Uris Hall, Ithaca, NY 14853-2501, USA
[b] Institute of Economic Policy, Leibniz University Hannover, Königsworther Platz 1, Hannover 30167, Germany

ARTICLE INFO

JEL:
C93
D83
J7

Keywords:
Discrimination
Gender
Race
Memory
Experiment

ABSTRACT

Remembering people is at the core of many social and economic relationships. We present evidence of systematic biases in the way we remember people, based on two experiments. The first experiment is conducted in a real professional setting - academia. Participants of two academic conferences are asked to recall 'who presented what' a month after attending the conferences. The second experiment is a controlled version of the first. Participants are shown pictures of people, matched with the title of a paper. We exogenously vary the relative shares of women and non-white individuals. In both experiments, we find evidence that women and ethnic minorities are more likely to be remembered in settings where they are in a small minority. In contrast, they are more likely to be confused with each other when they are in larger fraction. These findings are in line with a theory of categorization. People with minority attributes appear to be "blended together." We conjecture that these biases in remembering could have important implications for the formation of professional networks.

1. Introduction

Remembering people plays a key role in many social contexts and in the labor market, in particular. It is a necessary condition for forming social ties, and social ties are known to matter for labor market careers and performance. For example, it is well-known that social networks are a major channel used to find jobs (Calvó-Armengol & Jackson, 2004). However, there has been little research on whether there are systematic biases in memory along attributes such as gender or race. Such biases in remembering could lead to discrimination of certain groups (Belot, 2015). The key question we are interested in is the following: When we meet someone and later try to recall relevant information about that person, are we more likely to remember it correctly if they belong to a majority group? Or is it the reverse?

We know from research in psychology that race and gender are prime attributes encoded about others (Montepare & Opeyo, 2002; Shepherd et al., 1991; Tibbetts & Dale, 2007; Valentine, 1992). They are therefore distinctive attributes and are likely to help remembering minority individuals. On the other hand, these attributes may also lead to stereotyping and to the blending of people sharing these attributes. Fryer and Jackson (2008) propose a model of bounded memory where they conjecture that people are sorted into *categories*, defined as a vector of specific *attributes* (e.g. "white" & "woman"). People sorted into the same category are lumped together and cannot be distinguished from each other, and the number of categories is limited. In a world with limited cognitive resources, the optimal "categorization technology" lumps people with attributes that are encountered less frequently into broader categories. An important implication being that people from minority groups will then be sorted into coarser categories than people from the majority group, with whom interactions are more frequent. As a consequence, people sharing the same minority attributes are more likely to be confused with each other in settings where there is room for confusion (i.e., when there are a number of individuals sharing these same minority attributes).[1]

We conduct two studies, one in the field and one in a controlled computer-based environment. In both studies, our focus is on whether

☆ We thank seminar participants at the European University Institute, FAIR at NHH in Bergen, the University of Edinburgh, Cornell, Cologne, and University of Hannover and conference participants at the annual meeting of the German Economic Association, EALE, the annual meeting of the VfS and the GfeW. We are especially thankful to Tommaso Battistoni for excellent research assistance. Marina Schröder thanks the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 456592273 for financial support.

* Corresponding author.
E-mail address: schroeder@wipol.uni-hannover.de (M. Schröder).

[1] The model of Fryer and Jackson (2008) builds the foundation for the research in this project. Clearly, there are other biases that may lead to discrimination such as for example representativeness heuristics (Tversky and Kahnemann 1974) or associative memory (Enke et al., 2020).

https://doi.org/10.1016/j.socec.2023.102008
Received 3 March 2022; Received in revised form 10 March 2023; Accepted 12 March 2023
Available online 14 March 2023

people can remember relevant information about a person. In the field, participants from two conferences are asked to match pictures of presenters with titles of papers. They are offered four options as possible answers. As a second task, they are also asked to write the name and institution of a presenter shown on a picture. We replicate the first task in a controlled computer-based setting, where experimental participants first see a sequence of pictures of people matched with titles of papers and, in a second stage, are asked to match pictures with titles, again in a multiple choice setting.

Academia is an interesting environment because it shares many characteristics with other labor markets, such as the reliance on social networks for recruitment or activities relevant to promotions (e.g. recommendations) and concerns regarding the success of specific groups of the population such as women and non-whites. There is indeed ample evidence showing that women and non-whites fare less well in academia compared to peers from majority groups (Blackaby & Frank, 2000; Blackaby et al., 2005; Carter et al., 1999; Kahn, 1993). Notably, there is a belief that social networks may play an important role in explaining differentials according to gender and race; which led leading professional institutions to invest specifically in improving the social networking opportunities of these sub groups.[2] One notable advantage of academia relative to other labor markets is, as Blackaby et al. (2005) point out, that it has readily available measures of productivity that are comparable across institutions (such as number of publications, ranking of publications, etc.). That is, it is possible to compare researchers working in different institutions, which is usually very difficult to do in other labor markets.

The analysis of the field setting (where women and ethnic minorities are in a minority) shows that, controlling for the presenters' academic achievements, female presenters are more likely to be correctly matched to the title of their paper. But participants cannot recall their names or institutions better than they can for men. We do not find significant differences according to race.

As a second step, we exploit an important feature of the experimental design – the multiple choice format of the picture-title matching task. In these questions, the correct answer (presenter or title) must be identified among a set of four possible options. Due to the random nature of the alternative options, the choice set sometimes includes another presenter of the same gender or race as the "correct" presenter; sometimes it does not. Our prediction is that if minorities are sorted into broader categories, confusion is more likely to arise in the first case than in the second. In contrast, people who belong to the majority group (e.g. white men) should be less likely to be confused with others who share the same attributes. We find evidence that this is the case: people are more likely to confuse women with each other than they are to confuse men with each other.

The field setting has the limitation that the minorities correspond to a limited number of people, and it could be that our results are driven by idiosyncratic characteristics of these people. We therefore also provide evidence from a complementary controlled on-line computer-based experiment replicating key features of the field setting: In a first stage of the controlled on-line experiment, participants see pictures of people matched (at random) with titles of papers in economics. In a second stage, participants are asked to either match a person to one of four titles, or a title to one of four people. The people appearing on pictures are drawn from a picture database and are not real economists. The advantage of the controlled setting is that we can use a larger battery of people, implement a random assignment of titles to people, guarantee that participants have no other information about these people than the information we give them, and we can contrast treatments where there is a balance of people of both gender and race, with treatments where females or non-whites are in minority.

The results from the controlled setting confirm the findings from the field: We find evidence of a positive bias in remembering women when they are in minority. We also find that this positive bias completely disappears in the treatment where there is an equal proportion of men and women, as predicted. The pattern is less clear for non-white individuals.

As in the field setting, we also exploit the multiple choice format of questions to test for confusion. We find evidence that women and non-whites are more likely to be confused with others who share the same gender or race, particularly in settings where they are not in minority, that is, in contexts where there is more room for confusion.

To our knowledge, we are the first to provide evidence of biases in remembering information about individuals. We believe these findings have implications for our understanding of the structure of social networks and how they may impact success in real professional markets. In academia for example, it is clear that remembering plays an important role in network formation: We often rely on memory (who we remember meeting) to come up with names of potential candidates for seminars, workshops or even employment positions. The results of this study raise concerns about the practice of relying on memory for activities that may affect people's careers.

The rest of the paper is structured as follows. Section 1 presents the experimental design for the field setting and Section 2 presents the experimental evidence collected for the field setting. Section 3 presents the experimental design and evidence from the controlled setting. We conclude in Section 4.

2. Experimental design of experiment 1 (real professional setting)

We conducted an on-line experiment following two "plenary sessions" economics conferences. The first conference was the "Deception, Incentives and Behavior" conference, which took place in San Diego in April 2012 at the Rady School of Management at UCSD (in the following, we will refer to this conference as the San Diego conference). The second conference was the annual European "Search and Matching Conference" (organised by the professional network SaM – "Search and Matching"), which took place in Edinburgh in May 2014 (in the following, we will refer to this conference as the Edinburgh conference). Both conferences were single-session conferences, i.e. all talks were plenary talks. The presentations lasted for 10 or 20 min in San Diego and for 30 min in Edinburgh. One important advantage of the single session setup is that participants do not self-select across sessions, which would complicate the interpretation of the results if that were a possibility. Most respondents attended all sessions. Nevertheless, it is possible that some participants did not attend some of the presentations, and it is important for the analysis not to confuse participants who cannot remember a presenter because they did not attend or because of imperfect memory. We will come back to this issue in the analysis.

We contacted all participants a month after each conference and asked them to participate in our experiment. We offered $50 Amazon vouchers to the five best performers in the memory task (that is, 10 vouchers in total for both data collections). These were distributed by e-mail (participants did not need to be identified to receive the voucher).

The questionnaire consisted of three memory tasks. The time to answer the questions was restricted in each one of these tasks (see online Appendix for screenshots), such that it would be very difficult for participants to check information on the internet.

The first two tasks consisted of matching pictures of presenters to the title of the paper they presented. The pictures were obtained from public

[2] Prominent examples of such objectives are the professional network associations such as the "Black British Academics" or the "American Economic Association Committee on the Status of Women in the Economics Profession."

websites (such as homepages).[3] In the first task, participants saw *one paper title* and *four pictures* of conference participants.[4] The second task was the reverse of the first task. In the second task, participants saw only *one picture* of one presenter and *four titles* of papers.

Each of these tasks included six questions. Participants had 25 s to answer each of the questions. We set this time limit in order to avoid cheating (i.e., on-line search during the experiment). To determine the time, we asked student assistants who did not face any time restrictions and any performance contingent incentives to take part in the survey. Given that the student assistants did not attend the conferences, we also did so ourselves. We chose to set the time limit according to the maximum amount of time needed to answer a question in this pre-test. When selecting options for the first and second tasks, we made sure that we only asked one question about the same presenter. In these two tasks, participants know that one of the four options is the correct answer. In both cases, the options were randomly chosen from the pool of presenters and the pool of papers (from the same event). This feature is helpful because there will be a random variation in whether there are other people sharing attributes such as gender and race in the same choice set. Because minority groups are by definition small in size, most answer choice sets will have a majority of presenters from the majority groups and a minority of presenters from the minority groups. As a consequence, presenters from minority groups may be matched more accurately to the paper they presented, simply because of these distinctive characteristics. Indeed, if gender and race are prime attributes recorded in memory, participants may remember that the presenter had these attributes (even though they may not recall specifically who that presenter was). But we would also expect them to be more likely to be confused with someone sharing these attributes if confusion is possible (i.e., if there are more of them in the choice set).

Finally, the third task requires participants to provide the name and current academic institution of a presenter, shown on a picture. The answer here is free format (participants must write something and cannot choose between pre-specified answers) and allows us to evaluate the ability to recall the identity of a specific presenter. Here, remembering distinctive attributes is obviously not sufficient to provide a correct answer, and the task is harder because it is not a multiple choice format, such that getting it right by guessing is less likely. Again, participants had 25 s to answer each question and were asked to identify six presenters.

We should briefly justify the key design choices (matching of pictures with names and papers) we made. We could have chosen to ask about alternative pieces of information such as presentation slides, how well someone presented or the main finding of the presentation. The worry is that presentation slides could contain identifying information and that the perception of presentation quality or main findings is subjective and therefore difficult to control for. The advantage of a title is that it is a valuable piece of information, that is objective and does not contain identifying information. There are of course other pieces of information that are perhaps more important to remember than the title of a paper, such as how interesting the paper is or the general findings. The challenge is that if that information is not purely objective (as a title is), then it is difficult to operationalize in an experiment: we would not have a measure of how interesting a paper was for example. Similarly, general findings, or presentation slides, are likely to vary in quality in a way that we cannot fully control for. This means that we would not be able to know whether differences in recall are due to differences in the quality of the information presented, or due to differences in individual attributes of the presenter. The title on the other hand, is a neutral and objective information, and titles in a conference are presented in a systematic manner – on the first slide. This is why we chose to focus on titles. Also, we focus on matching pictures of presenters with relevant information (names or papers) because this is what is most relevant in social networking settings. The match between names and papers is also important in academia, but in contrast to faces of presenters, it is easy to retrieve a name of an author of a specific paper just by searching on-line. The reverse is also true. The face of someone on the other hand, is not searchable (or at least not yet).

After completing the three tasks, we asked respondents to answer a short survey. We asked about demographics of the respondents such as gender, age, field, research position, and race. We asked them to indicate which sessions they attended, and for the Edinburgh conference, we additionally asked respondents to indicate whether they knew the presenter before the conference. For the analysis, we will only consider responses to questions involving presenters in sessions that respondents report having attended. This is to minimize a potential selection bias and to ensure that the reason why participants do not recall a presenter is because of imperfect memory and not because they did not attend.

In both experiments, those participants who wanted to participate in the contest for an Amazon voucher were asked to provide an e-mail address. We had multiple versions of the questionnaire for each event (four for each conference), varying the presenters involved. Our data includes questions about 37 of the 44 presenters from the San Diego conference[5] and all 25 presenters from the Edinburgh conference. Not all presenters are featured in each task though. The allocation of presenters to tasks was random.[6]

Participants earned points for each correct answer. They earned two points per correct answer for the first two tasks. For the third task, they earned one point for a correct name and one point for a correctly indicated affiliation.[7] Vouchers were awarded to those respondents who achieved the highest number of overall points for the three tasks.

Finally, we collected background information of presenters by searching information on-line. We recorded information about each presenter's current academic position, number of publications, number of top five journal publications, current research institution, and if applicable, time since completion of the PhD. We collected most of this information from personal webpages and CVs. We also recorded the ranking of the institution the presenter was affiliated to at the time of presentation, using the Tilburg Economics Schools and Research Ranking (https://econtop.uvt.nl/). We constructed a measure for race using the on-line service kairos (https://www.kairos.com/diversity-recognition).[8] The software identifies the racial background of a person based on a picture and indicates relative percentages distinguishing between five categories (e.g. 0% black, 90% white, 5% Asian, 5% Hispanic, 0% other). Note, that this categorization based on pictures is only a proxy for actual race. In our analysis, we focus only on whether an individual is from a racial minority or not. Thus, we categorized a

[3] One concern that immediately arises is that there may be biases according to dimensions such as gender and race in how accurate the picture is. We are able to address this concern by asking independent raters to evaluate the similarity of the public picture with a picture taken in the conference room itself (for the Edinburgh conference). See on-line appendix for results with an additional control for the similarity of pictures.

[4] The set of pictures mixed presenters and other attendants of the conference who presented a poster.

[5] We could only include those with pictures on their public website and excluded one presenter from the San Diego conference because the presentation was not of an academic paper. Among the other presenters, we selected randomly.

[6] As a consequence, it is possible that participants to the two conferences were asked to identify themselves. This should not bias the results in any specific way though, but could raise the level of accuracy slightly.

[7] When assigning points and for the following analysis, we allowed for spelling mistakes and abbreviations.

[8] At the time, the software was available on-line and was free to use, it is no longer the case. In addition to the categorization through the software, we asked student assistants to do the categorization. Both categorizations deviate only in one case and results are robust when using this alternative categorization through student assistants (see Tables A1 and A2 in the on-line Appendix).

presenter as non-white if the dominant category was not white and as white if it the dominant category was white. Due to the small number of non-white presenters, we will not further distinguish between different racial minorities.

3. Results (Experiment 1)

The goal of the analysis is to establish whether there are systematic biases in memory according to gender and race, controlling for productivity variables. To test for systematic biases in remembering, we study the accuracy of memory conditioning on a large set of variables that are likely to correlate with "academic productivity", such as the seniority of the presenter, the number of publications in top ranked peer reviewed journals and the rank of the current institution of the presenter. Of course, some of these variables are likely to correlate with how known a person is and on past interactions as well. These variables could also directly correlate with the quality of the presentation itself, which may make it more memorable. We will not be able to tease out between these alternative explanations. What matters though for our research question is whether there are systematic biases in memory along variables such as gender and race, conditioning on variables capturing productivity.

We first present summary statistics on presenters and respondents (3.1). We then study how the presenters' characteristics relate to the accuracy of recall (3.2) and test more specifically for the theory of categorization (3.3).

3.1. Summary statistics (Experiment 1)

Table 1 presents summary statistics of the presenters included in the experimental study by conference and seminar series. There were 44 presenters in San Diego, of which 37 are included in the experiment, and all 25 presenters from the Edinburgh conference are included.

The most notable differences between the events are in the percentages of women (about one third in the San Diego conference and one fifth in the Edinburgh conference) and the percentages of economists (the San Diego conference was interdisciplinary, with a majority of economists, while the Edinburgh conference was dominated by economists).

Table 2 presents summary statistics of the respondents, again split by event. There were 114 participants in the San Diego conference, among which 41 participated to the experiment; and 111 in Edinburgh, among which 46 participated.

Table 1
Summary statistics of the presenters included in our study (Experiment 1).

Characteristics of presenters means (standard deviations)	San Diego April 2012	Edinburgh May 2014
# of presenters	37	25
% female	35%	20%
% non-white	11%	16%
% non-native English speakers	73%	64%
# of years since PhD	12.6 (9.8)	13.9 (11.9)
rank current institution	51.1 (38.1)	59.36 (37.54)
# top 5 publications in economics	1.6 (3.0)	2.3 (4.1)
# of publications	26.1 (27.3)	17.8 (22.5)
% economists	73%	96%

Table 2
Summary statistics of the respondents (Experiment 1).

Characteristics of respondents means (standard deviations)	San Diego	Edinburgh
# of respondents	41	46
% female	36%	30%
% tenured	33%	33%
% non-white	24%	20%
% economist	74%	100%
% over 40 years old	26%	26%

To ensure anonymity, the information we collected about respondents is more limited and coarse. About a third of our respondents are female and only a quarter above 40 years old. Except for the fact that all respondents are economists in the Edinburgh conference, against three-quarters at the San Diego conference, there are no large differences in the respondents' characteristics across the two conferences.[9]

Table 3 presents summary statistics regarding the performance in each task. Overall, we find that people are better able to map faces to paper titles than recalling names or institutions. Of course and as mentioned earlier, when mapping faces to titles, the participants had a multiple-choice question, so guessing was easier. The mean accuracy rate in mapping faces and titles is around 60% on average for the two conferences. The accuracy is lower for task three (recalling name and institution of a person) at around 40%.

One important issue we wish to address head on is selection. We have two potential sources of selection: One is in responding the survey, the second is in attending the presentation or not.

Regarding the first source of selection, in all events, we have more than a third of attendees responding to our survey, and the invitation e-mail did not give any hint as to what the survey would involve, except that it was related to the event of interest (conference attended). This response rate is in fact quite high relative to typical response rates in social sciences surveys. Cook et al. (2007) report that one should expect between a 25% and 30% response rate from an e-mail survey when no follow-up takes place. Of course, it would have been helpful to be able to compare characteristics of respondents and attendees, but we do not have the information on attendees. As is usually the case with surveys, we only have information on those who actually responded. But we have a priori no reason to believe that those who self-selected into responding would exhibit stronger gender or racial biases in memory than those who did not respond.

The second issue is key for the interpretation of the results. People could fail to recall someone either because of imperfect memory (what we are interested in) or because they did not attend the presentation. If participants are less likely to attend sessions of presenters with certain

Table 3
Summary statistics of performance in each task (Experiment 1).

	San Diego	Edinburgh
tasks 1 and 2 (picture and title)	.57	.65
	(0.50)	(0.48)
task 3 (picture and name/institution)	.43	.38
	(0.50)	(0.49)

Accuracy rate: Mean values with standard deviations in parenthesis.

[9] Since we collected information on the respondents as well, we are able to study to what extent their characteristics correlate with accuracy of recall (see Table A3 in the on-line Appendix). While women perform slightly better and non-whites slightly less well in this task, we do not find evidence of any significant variable affecting accuracy of recall, except for the respondent being an economist.

characteristics (minorities for example), we would have a selection bias if we would categorize these people in the same category as those who cannot recall the presenters with these characteristics conditionally on attending. The ideal setup for our research question is one where either attendance is random or attendance is compulsory for everyone. As already mentioned, we chose these particular events precisely because we expected little selection into attendance to take place. In contrast to larger conferences, the norm in these smaller one-session events is very much that everyone attends all sessions, although of course in practice some people do not attend some of the sessions.

Table 4 presents summary statistics about self-reported attendance at both conferences. We have a mean attendance of 85% for all sessions at the San Diego conference (with a median of 86%) and a mean attendance of 82% at the Edinburgh conference (with a median of 85%). These high attendance rates should reassure that self-selection into attendance is not a major issue. Further, we find no evidence that attendance is different for women or non-whites (see Table A4 in the online Appendix). To minimize any chance of bias, we will limit the analysis to those who reported they attended the session and we will check whether attendance is correlated with presenter characteristics.

Table 4
Summary statistics on self-reported attendance at each session (Experiment 1).

	San Diego	Edinburgh
Average	0.85	0.82
Median	0.86	0.85
Max	1.00	1.00
Standard deviation	0.11	0.09
Min	0.57	0.70

3.2. Presenter characteristics and accuracy of recall (Experiment 1)

We pool the data from both conferences for the analysis, and we also pool the data from the first two tasks, as they both involve matching pictures of people to titles of papers.[10] In Tables 5 and 6, we present the results of a two-way error component linear probability model, allowing for presenter and respondents' random effects.[11] The dependent variable is a dummy equal to 1 if the answer was correct and equal to 0 otherwise. Note that we only consider answers of respondents who also attended the corresponding session.

Table 5 reports the results related to the task of mapping pictures of presenters to titles of the paper they presented (multiple choice type questions). Column (1) shows the estimates of a model conditioning on gender and race.[12] Column (2) conditions in addition on characteristics correlated with the productivity or expertise of the presenter. We control for rank of current institution, number of publications, number of top 5 publications in economics, a dummy for being an economist, number of years since PhD completion, and a dummy for presenter being a native English speaker (we conjecture that being a native English speaker may be correlated with the quality of the presentation). Column (3) additionally controls for characteristics of social proximity. We have three main variables of social proximity: *Same field as respondent* (which is a dummy equal to 1 if the presenter and respondent are in the same field and equal to 0 otherwise), *same gender as respondent* (which is a dummy equal to 1 if both the presenter and respondent are of the same gender, and equal to 0 otherwise), and *same ethnicity as respondent* (which is a dummy equal to 1 if both are white and equal to 0 otherwise).[13]

We find that participants are much more likely to accurately match a female presenter with the title of the paper presented, compared to a male presenter. The estimated difference is 14 percentage points, once we control for academic achievements and social proximity (Column (3)). We find an effect of similar magnitude for non-white, that is significant at the 10 percent level. It is notable that the gender and non-white dummies becomes larger and significant (at the 5 percent and the 10 percent level, respectively) when adding controls for productivity. This suggests that women and non-whites are on average less established than men and whites in our sample are and it is therefore unlikely that unmeasured productivity-related variables would explain the positive coefficients.

Academic achievements do matter as well, but not to a large extent. All else equal, one needs nine top 5 publications in economics to achieve a similar improvement in recall as being female or non-white. This is

Table 5
Probability of correct mapping between face and title (Experiment 1).

	(1)	(2)	(3)
Female	0.082	0.119**	0.135**
	(0.056)	(0.059)	(0.058)
Non-White	0.074	0.103	0.135*
	(0.075)	(0.075)	(0.076)
Non-native English speaker		0.056	0.060
		(0.059)	(0.057)
# of years since PhD		0.005	0.004
		(0.004)	(0.004)
Rank current institution		−0.000	−0.000
		(0.001)	(0.001)
# of top 5 publications in economics		0.016*	0.015*
		(0.009)	(0.008)
# of publications		−0.000	−0.000
		(0.001)	(0.001)
Economist		0.018	−0.173**
		(0.078)	(0.088)
Same field as resp			0.252***
			(0.061)
Same gender as resp			0.038
			(0.034)
Same race as resp. (both white)			0.055
			(0.045)
Constant	0.607***	0.464***	0.359***
	(0.034)	(0.092)	(0.097)
Observations			
	892	850	850
Number of groups	1	1	1

Two-way error component linear probability model, allowing for presenter and respondents' random effects. The dependent variable is a dummy for correct answer. Independent variables: female (binary, equal to 1 for female presenters and 0 otherwise), non-white (binary, equal to 1 for non-white presenters and 0 otherwise), non-native English speaker (binary, equal to 1 for presenters who are not native in English and 0 for presenters who are native speakers in English), # of years since PhD (difference between the year the presenter obtained the PhD and the year the conference took place), rank current institution (rank of the current institution of the presenter based on Tilburg Economics Schools and Research Ranking), # of top 5 publications in economics (sum of presenter publications in the American Economic Review, Econometrica, the Journal of Political Economy, the Quarterly Journal of Economics, and the Review of Economic Studies), # of publications (sum of peer-reviewed publications of the presenter), economist (binary variable equal to 1 if presenter is an economist and 0 otherwise), same field as resp. (binary, equal to 1 if presenter and respondent are in the same field and 0 otherwise), same gender as resp. (binary, equal to 1 if the presenter and the respondent are either both female or both male and equal to 0 otherwise), same race as resp. (binary, equal to 1 if respondent and presenter are white and 0 otherwise).
Standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

[10] We replicate the analysis for each conference separately (see Table A5 in the on-line Appendix). We should however take the results of this analysis with caution because the number of presenters falling into the different categories of interest (female and non-white) is small when considering each conference separately. Overall, the effects differ somewhat across events, but they are not inconsistent with each other.

[11] The model is estimated in STATA 16.0 using the command "xtmixed".

[12] We do not control for the interaction of female and non-white. Of the non-whites in our sample, 37 percent are female (50 percent in the San Diego conference and 25 percent in the Edinburgh conference).

[13] Due to the minority status, we cannot control for settings in which respondents and presenters are of the same race but not white.

Table 6
Probability of correctly remembering the name and institution (Experiment 1).

	Correct answer name			Correct answer institution		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−0.097	−0.015	0.019	−0.030	0.034	0.058
	(0.086)	(0.061)	(0.066)	(0.078)	(0.063)	(0.066)
Non-white	−0.019	0.107	0.146*	0.020	0.084	0.137*
	(0.096)	(0.069)	(0.080)	(0.087)	(0.071)	(0.080)
Non-native English speaker		−0.100*	−0.084		−0.077	−0.058
		(0.060)	(0.064)		(0.062)	(0.064)
# of years since PhD		0.006	0.006		0.000	−0.001
		(0.004)	(0.004)		(0.004)	(0.004)
Rank current institution		−0.001*	−0.001		−0.002***	−0.002**
		(0.001)	(0.001)		(0.001)	(0.001)
# of top 5 publications in economics		0.015	0.015		0.016	0.015
		(0.010)	(0.011)		(0.011)	(0.011)
# of publications		0.002	0.003*		0.002	0.003
		(0.002)	(0.002)		(0.002)	(0.002)
Economist		0.048	−0.230**		0.096	−0.157
		(0.081)	(0.109)		(0.084)	(0.111)
Same field as resp			0.343***			0.311***
			(0.081)			(0.084)
Same gender as resp			0.070			0.045
			(0.045)			(0.046)
Same race as resp. (both white)			0.088			0.112*
			(0.058)			(0.060)
Constant	0.463***	0.360***	0.198*	0.427***	0.412***	0.256**
	(0.050)	(0.091)	(0.111)	(0.047)	(0.095)	(0.111)
Observations	429	429	429	429	429	429
Number of groups	1	1	1	1	1	1

Two-way error component linear probability model, allowing for presenter and respondents' random effects. The dependent variable is a dummy for correct answer. Independent variables: female (binary, equal to 1 for female presenters and 0 otherwise), non-white (binary, equal to 1 for non-white presenters and 0 otherwise), non-native English speaker (binary, equal to 1 for presenters who are not native in English and 0 for presenters who are native speakers in English), # of years since PhD (difference between the year the presenter obtained the PhD and the year the conference took place), rank current institution (rank of the current institution of the presenter based on Tilburg Economics Schools and Research Ranking), # of top 5 publications in economics (sum of presenter publications in the American Economic Review, Econometrica, the Journal of Political Economy, the Quarterly Journal of Economics, and the Review of Economic Studies), # of publications (sum of peer-reviewed publications of the presenter), economist (binary variable equal to 1 if presenter is an economist and 0 otherwise), same field as resp. (binary, equal to 1 if presenter and respondent are in the same field and 0 otherwise), same gender as resp. (binary, equal to 1 if the presenter and the respondent are either both female or both male and equal to 0 otherwise), same race as resp. (binary, equal to 1 if respondent and presenter are white and 0 otherwise).
Standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

perhaps surprising, as one would expect well-established researchers to be more likely to be remembered for several reasons: The value of social ties may be higher, they are likely to give higher quality presentations, and they are more likely to have been previously encountered. But we find that these factors do not matter much for recall accuracy. Being in the same field, on the other hand, is a strong and significant (at the 1 percent level) predictor of recall accuracy. On top of that, all else equal, information about economist presenters are less likely to be remembered. This is consistent with a value of standing out – economist respondents (who are in majority) can better distinguish between presenters who are psychologists in comparison to psychologist respondents, who have a harder time distinguishing between the large fraction of presenters who are economists.

Table 6 shows the results for the task consisting of providing the names of presenters and of their institutions (Task 3). Columns (1) to (3) relate to the naming of presenters and Columns (4) to (6) relate to the naming of institutions. Here, remembering of distinct attributes alone is not sufficient to get a correct answer. As in the previous Table, Column (1) includes gender and race dummies; Column (2) controls for productivity and expertise, and Column (3) controls for social proximity variables.

Here we find evidence of a minority bias in favour of non-whites. Measures of academic performance and establishment do matter somewhat, but their effects are very small in magnitude. We find that being in the same field as the presenter matters greatly and in a similar way as for mapping between faces and titles. Of course, one obvious explanation for the effects of social proximity and of academic achievements is that the respondent is more likely to know the presenter if they are in the same field and if the presenter is well established. This is a compelling story, but one we should be very cautious with, as knowing someone is not exogenous and is likely to be correlated with factors affecting how memorable someone is. There is an obvious circularity in the fact that respondents can only claim to know someone if they remember that person. Nevertheless, we collected information on whether participants claimed to know the presenter for the second of the two conferences (Edinburgh). We find that knowing the presenter is a strong predictor of accuracy of recall in all tasks. The effects of being female or non-white are similar in magnitude but not statistically significant in Edinburgh (whether one controls for knowing the presenter or not). The size of the coefficients remains similar to those excluding the dummy variable for "knowing the presenter" (see Table A6 in the online Appendix for detailed results).

Turning to the naming of institutions (Columns (4)-(6)), we do not find evidence of significant biases in the accuracy of recall according to gender. We find that respondents perform slightly better at naming the institutions of non-white presenters. People are better able to remember the institution of the presenter if the institution is highly ranked, although again the magnitude of the effect remains modest. We find evidence of a social proximity effect. We find that being in the same field matters. Additionally, we find that there is a significant (at the 10 percent level) same race effect, where white respondents are better at remembering white presenters. One has to be careful when interpreting this result, because due to the minority status of non-whites in the sample, the control for same race as respondent only includes observations where both the presenter and the respondent are white but no observations where they are both of the same race but non-white.

3.3. Probability of confusing people with similar attributes (Experiment 1)

Our findings so far show that information about women is more likely to be remembered accurately, in a setting where they are in minority. However, this is only the case in the multiple-choice type questions, where due to the minority status of women, remembering the gender may be sufficient in order to correctly identify a female presenter. In a next step, we evaluate whether minorities are more likely to be confused with those who share these minority attributes. This question directly relates to the theory of categorization.

For this analysis, we use data from tasks 1 and 2. In those two tasks, the respondent is asked to pick between four pictures of presenters and correctly match it to a presented paper (task 1) or to pick between four titles and correctly match it to a picture of a presenter (task 2). Because in each task, we chose the other three options at random, presenters from minority groups will often be shown with other presenters who are from majority groups. But, the design also allows us to study whether they are more likely to be confused with someone who has the same gender or race.

We test whether minority groups are more likely to be confused when there are several of them as possible answer options. To do this, we estimate a conditional logit model, where the dependent variable is a dummy indicating whether or not an option is chosen, and the independent variables are attributes of the various options. The advantage of the conditional logit model is that it only uses variation in attributes within each choice set (i.e., between the four options) to estimate the effects of these attributes on the probability of being chosen.

We also control for a variable called "same session", which is a dummy equal to one if the presenter shown as one of the four options presented in the same session as the presenter corresponding to the correct answer and zero otherwise. We can consider "same session" as a proxy for similarity in research topic, as the sessions in the Edinburgh conference were organised according to topics, in the San Diego conference all talks were on one topic though (deception).

The results of the conditional logit model are reported in Table 7. We find that respondents are significantly more likely to choose the correct option compared to the other options, but we do not find evidence for confusion (i.e., choosing another option of the same gender and race instead of the correct option) overall. The theory of categorization proposed by Fryer and Jackson (2008) would predict that confusion should be more likely to occur for minorities and/or for people who have attributes associated with a lower expected value of interaction.

Table 7
Conditional logit for probability of correct mapping between face and title – distinguishing by the attributes of the correct answer (Experiment 1).

	(1) all	(2) Male Presenter	(3) Female presenter	(4) White presenter	(5) Non-white presenter
Correct option	0.363***	0.386***	0.258*	0.359***	0.364*
	(0.048)	(0.049)	(0.147)	(0.049)	(0.216)
Option same gender	0.005	−0.035	0.084**	−0.017	0.216***
	(0.026)	(0.031)	(0.041)	(0.028)	(0.064)
Option same race	0.045	−0.023	0.180***	0.021	0.176
	(0.029)	(0.036)	(0.032)	(0.032)	(0.110)
Presented in same session	0.117***	0.125***	0.053	0.141***	−0.160
	(0.043)	(0.046)	(0.127)	(0.045)	(0.179)
Observations	4196	3132	1064	3788	408

The dependent variable is a dummy indicating whether an option was chosen. Independent variables: correct option (binary, equal to 1 for the correct option and 0 for incorrect options), option same gender (binary, equal to 1 if option is not the correct but individual corresponding to this option is of the same gender as the individual corresponding to the correct option and zero otherwise), option same race (binary, equal to 1 if option is not the correct but individual corresponding to this option is of the same race as the individual corresponding to the correct option and zero otherwise), presented in same session (equal to 1 if option is not correct but presenters corresponding to the option presented in the same session as the presenter corresponding to the correct option).
We report average marginal effects.
Standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

In Table 7, Columns (2) to (5), we estimate the conditional logit for different subsamples, corresponding to the gender and race of the speaker associated with the correct answer. We find that men are not more likely to be confused with other men. Women, in contrast, are significantly (at the 10 percent level) more likely to be confused with other women. Interestingly, race also seems to play an important role when it comes to remembering women, while this is not the case when remembering men. For non-whites, we find evidence that gender is used as an attribute (more confusion is likely with a person from the same gender). Again, we do not find similar effects when it comes to remembering the majority group, i.e., white individuals. Our results are in line with the theory of categorization discussed earlier, whereby broader categories are formed, for women in particular.

One drawback of the field is that minority groups are by definition in minority, so our results could be driven by idiosyncratic characteristics of those with minority attributes; and we have limited scope to test for the theory of categorization. Ideally, we would like to have a larger set of people with minority attributes and vary exogenously the relative sizes of groups with minority attributes. As discussed earlier, we predict that minorities may be more likely to be confused with each other when there is more room for confusion, that is, when their share is relatively large. We conduct a controlled on-line experiment that will allow us to address the caveats of the field setting.

4. Experiment 2: controlled setting

The second experiment is also conducted on-line but is not based on a real setting. The experiment replicates key features of the field setting. We matched pictures of people drawn from a picture database[14] to titles of NBER working papers (drawn at random from the NBER paper series in 2019). The picture database has pictures from individuals of different gender and race; we used those from the categories "Caucasian white" and "Asians". In the latter, we selected pictures of East Asians and excluded pictures of South Asians. Thus, in the second experiment, we focus only on East Asians when studying racial minorities. We decided to conduct the experiment using East Asians, as East Asians are well represented in the subject pool that we used to conduct this experiment.

In a first stage, participants see 12 pictures of people randomly matched with titles of papers. Each match is displayed for 10 s. In a second stage and as in the first experiment, participants are asked to either match a person with a title (choosing between four possible titles), or to match a title with a person (choosing between four possible people). Participants had 10 s to answer each of the questions.[16] Four questions were asked in total, one with the correct answer being a white man, a second a white woman, a third an East Asian man and a fourth an

[14] Stimulus images courtesy of Michael J. Tarr, Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University, http://www.tarrlab.org/. Funding provided by NSF award 0339122. We used a subset of pictures (we removed those where the person had uncommon accessories such as a hat or a piercing), consisting of 75 pictures of Caucasian whites and 29 East Asians. To avoid the task to be about picture recognition rather than person recognition, we use pictures with different profile orientation (front facing or face turned left or right by 15 to 45 degrees).

[15] The distribution of respondent characteristics does not differ significantly across treatments except for one case: There are significantly more East Asian respondents in the "Female & non-white minority" treatment than in the Balanced treatment (p-value of Fisher test, p=0.062). Robustness checks do not indicate that this imbalance drives any of the results.

[16] To set the time frames, we again asked student assistants who did not receive any performance contingent incentives to participate in the survey without time limitations. We then set time limits corresponding to the maximum amount of time needed in this pre-test.

East Asian woman (the order was randomized). Two of the three incorrect answers corresponded to titles or people shown in the first stage (chosen at random), one was a completely new person or title, drawn at random, but with the constraint that the race and gender of the new option was with equal chances a white man, white woman, East Asian man or East Asian woman. This clustered randomization ensures we have a variation in the characteristics of the choice set and allows us to examine again, whether people just remember attributes (gender, race) of the correct answer or are able to remember who the specific person was.

We implemented four treatments as summarized in Table 8. We recruited 387 participants from the subject pool of the experimental laboratory at the University of Cologne. The summary statistics of these participants are shown in Table 9 below.[17] In all treatments, we have a majority of white and female respondents. The gender composition is however more balanced than the racial composition.

Table 10 reports the mean accuracy rates in remembering, according to the attributes of the person corresponding to the correct answer and the treatment. These descriptive statistics point at a positive minority bias especially in the female and East Asian minority treatments. Comparing the recall accuracy of different groups in the balanced treatment, we find no significant differences.

In Table 11, we present estimates of the probability of a correct answer, depending on the gender and race of the person corresponding to the correct answer. In line with our findings from the field, we find that the probability that a person is matched correctly to the title is higher when the correct answer involves a woman. However, this difference arises only in the treatments where they are in minority. When women are a minority, they are 15 to 18 percentage points more likely to be correctly matched to the title of the paper they were matched with in the first stage. We provide evidence of a bias in favour of East Asians only in the treatment where both women and East Asians are in minority. Whether the respondent shares the same gender or race with the person corresponding to the correct answer does not matter except in the balanced treatment, where having the same gender enhances recall.[18]

In Table 12, we distinguish specifically between settings in which women or East Asians are in minority and those in which they are not. For example, white women are in minority in two treatments ('Female minority' and 'Female & East Asian minority') and are not a minority in two others ('Balanced' and 'East Asian minority'). We report the results for the relevant treatments in which the relevant group is not a minority in Panel A and those in which the relevant group is a minority in Panel B of Table 12. Since white men are never in the minority in any treatment, we report results only for white men in Panel A.

We find evidence for differences in the way women and East Asian are remembered. However, we find such differences only when they are not in minority. For white men, we find no indication that people are more likely to confuse them with other white men. The options that are either white or male are not significantly more likely to be chosen. However, when the correct answer involves a white woman or an East Asian man, we find evidence of reliance on gender and/or race to select an answer. People are more likely to select a wrong option of the same gender and race. This is true for white women, and this is true for East Asian men. The evidence is weaker for the case where the correct answer corresponds to an East Asian woman, but there is only one treatment where they are not in minority (balanced), and so the number of observations is smaller. The point estimates are however quite large and thus we cannot reject a possible categorization effect.[19]

The analysis presented in Table 12 pools all the respondents, with no distinction according to their gender or race. To examine whether there are own-group biases in memory, we estimate the models in Table 12 separately for (1) white respondents only (our sample of East Asian respondents is too small to analyse separately), (2) women only and (3) men only (see Tables A11, A12, and A13 in the on-line Appendix). We find stronger evidence for a categorization according to gender if we restrict the sample to white respondents only. Regarding gender, we find evidence that the confusion effect is stronger for male respondents; we do not find significant evidence that female respondents confuse women with other women, while male respondents do.

Table 8
Treatments (Experiment 2).

	Distribution of pictures (numbers in each treatment)				Number of subjects
	White men	White women	East Asian men	East Asian women	
1 Balanced	3	3	3	3	99
2 Female minority	5	1	5	1	97
3 East Asian minority	5	5	1	1	96
4 Female and East Asian minority	9	1	1	1	96

Table 9
Summary statistics of respondents (Experiment 2).[15]

	Balanced	Female minority	East Asian minority	Female and East Asian minority
# of respondents	99	97	96	96
% female	59%	54%	55%	64%
% non-white	15%	16%	20%	24%
% East Asian	4%	6%	5%	11%

Table 10
Mean rate of accuracy in recall across treatments (Experiment 2).

Race and gender of "correct answer"	Balanced	Female minority	East Asian minority	Female and East Asian minority
White men	0.60 (0.49)	0.55 (0.50)$^{n.s.}$	0.64 (0.48)$^{n.s.}$	0.49 (0.50)$^{n.s.}$
White women	0.66 (0.48)	0.71 (0.46)$^{n.s.}$	0.70 (0.46)$^{n.s.}$	0.78 (0.42)*
East Asian men	0.60 (0.49)	0.54 (0.50)$^{n.s.}$	0.64 (0.48)$^{n.s.}$	0.73 (0.45)*
East Asian women	0.57 (0.50)	0.72 (0.45)**	0.66 (0.48)$^{n.s.}$	0.76 (0.43)***

Upper cases refer to results from a Fisher's exact test comparing results to the Balanced treatment.
Standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$; $^{n.s}p > 0.10$.

[17] We also study to what extent respondent characteristics correlate with accuracy of recall (see Table A3 in the on-line Appendix). We find that female respondents perform significantly better and non-white respondents significantly less well in this task.

[18] In Table A8 in the appendix, we provide a similar analysis including the interaction of gender and ethnicity. We find that the gender effects arise both for white and for East Asian females. We also replicate the findings of Table 10 including controls for gender and ethnicity of the respondents (see Table A9 in the Appendix).

[19] One may argue that these results are driven by the fact that there are no similar individuals to confuse minorities with. To test for such an alternative interpretation, we conducted a similar analysis restricting the data to those questions where there was an answer option of the same gender and the same race as the correct option. Results are robust to this alternative approach. For details see Table A10 in the on-line Appendix.

Table 11
Probability of correct mapping between face and title (Experiment 2).

	Balanced		Female minority		East Asian minority		Female and East Asian minority	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.015	−0.011	0.175***	0.177***	0.042	0.039	0.164***	0.148***
	(0.044)	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)	(0.043)	(0.046)
East Asian	−0.045	−0.107	−0.000	0.033	−0.023	−0.058	0.108**	0.156***
	(0.044)	(0.079)	(0.045)	(0.074)	(0.045)	(0.070)	(0.043)	(0.058)
Same gender as resp		0.130***		−0.021		0.033		0.055
		(0.045)		(0.046)		(0.043)		(0.045)
Same race as resp.		−0.076		0.043		−0.047*		0.076
		(0.082)		(0.075)		(0.072)		(0.058)
Constant	0.619***	0.634***	0.541***	0.515***	0.647***	0.670***	0.554***	0.478***
	(0.044)	(0.083)	(0.043)	(0.078)	(0.045)	(0.075)	(0.042)	(0.062)
Observations	396	396	388	388	384	384	384	384

Two-way error component linear probability model, allowing for picture and respondents' random effects. The dependent variable is a dummy for correct answer. Independent variables: female (binary, equal to 1 for female presenters and 0 otherwise), East Asian (binary, equal to 1 for East Asian presenters and 0 otherwise), same gender as resp. (binary, equal to 1 if the presenter and the respondent are either both female or both male and equal to 0 otherwise), same race as resp. (binary, equal to 1 if the presenter and the respondent are of the same race and equal to 0 otherwise).
Standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

Table 12
Conditional logit for mapping between face and title – by type of individual (Experiment 2).

Panel A - Treatments in which individual is not a minority

	(1)	(2)	(3)	(4)
Gender & race	White	White	East Asian	East Asian
correct answer	men	women	men	women
Treatments	1,2,3,4	1, 3	1, 2	1
Correct option	0.352***	0.422***	0.385***	0.354***
	(0.036)	(0.055)	(0.030)	(0.048)
Option same gender	0.048	0.134*	0.188***	0.125
	(0.051)	(0.072)	(0.029)	(0.076)
Option same race	0.081	−0.037	0.113***	0.127*
	(0.050)	(0.082)	(0.038)	(0.076)
Observations	800	384**	412	192

Panel B - Treatments in which individual is a minority

	(1)	(2)	(3)	(4)
Gender & Race	White	White	East Asian	East Asian
correct answer	men	women	men	women
Treatments		2, 4	3, 4	2,3,4
Correct option		0.358***	0.338***	0.460***
		(0.082)	(0.073)	(0.037)
Option same gender		−0.234	−0.121	0.039
		(0.149)	(0.084)	(0.078)
Option same race		−0.039	0.120	0.093
		(0.092)	(0.097)	(0.075)
Observations		380	340	532

The dependent variable is a dummy indicating whether an option was chosen. Independent variables: correct option (binary, equal to 1 for the correct option and 0 for incorrect options), option same gender (binary, equal to 1 if option is not the correct but picture corresponding to this option is of the same gender as the picture corresponding to the correct option and zero otherwise), option same race (binary, equal to 1 if option is not the correct but picture corresponding to this option is of the same race as the picture corresponding to the correct option and zero otherwise.
We report average marginal effects.
Standard errors in parentheses. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

5. Conclusion

This paper investigates the presence of systematic biases in memory in two different settings. First, we collected data in a real professional context, among researchers in Economics in the context of two international high profile conferences. We study how accurately conference participants can remember who presented what, a month after the conference. Specifically, conference participants were asked to match pictures of people to titles of papers presented and were asked to provide the presenters' name and institution based on their picture. We complement the evidence from the field with evidence from a controlled experiment where the individuals are matched at random to titles of papers and where we vary exogenously the fractions of men/women and non-whites.

We find evidence that people are better able to recall information about a person when the person involved is a woman and, to a lesser extent, a racial minority. We also show that women and racial minorities are more likely to be confused with others who share the same attributes. That indicates that minorities seem to be lumped into broad categories according to gender and race. These findings are in line with the theory of categorization proposed by Fryer and Jackson (2008) and with a distinctiveness effect that has been identified in previous studies on memory (Meissner & Brigham, 2001; Slone et al., 2000). People with minority attributes appear to be categorised according to these attributes and are "blended together". In settings where there are few people with minority attributes, recall is enhanced. In settings where there are more of them, this leads to confusion.

Both experiments presented in this paper have their strengths and their weaknesses. A central strength of experiment 1 is that it provides insights from a real professional context. Academic conferences share many attributes with other professional contexts. Remembering people that you briefly met at an organized event may for example also be relevant when it comes to business fairs, job fairs, or more generally networking events. A key challenge to study biases in such environments is that productivity is not readily observable. The benefit of the context of academia is that there are relatively good productivity measures. The challenge of the field setting is that almost by definition, presenters from minority groups are in a minority, and therefore the analysis of how minorities are treated is in fact relying on a small number individuals who belong to these minorities. In experiment 2, we abstract from a real professional setting with the advantages that we can increase the sample of individuals from minority groups and we can externally vary whether or not certain groups are a minority. The benefit of experiment 2 is that we can overcome central challenges of experiment 1 and gain insights on the mechanisms driving results. The downside of experiment 2 lies in the abstract character of the setting. We see these two studies as complementary to each other. In both designs, we concentrate on remembering of titles of presentations. In practice, pieces of information such as for example the quality of the presentation, the methods applied, or the general findings are also relevant. In our context, it was challenging to obtain good and unbiased measures of the quality of presentations for example. We controlled for variables measuring the overall academic quality of presenters, but were not able to study in a clean manner to what extent the quality of presentations affected recall. To do this, one would need to gather objective and unbiased measures of performance,

which is challenging. We leave this for further research.

Overall, we argue that biases in recall is an understudied but important research avenue; as such biases may have implications for career prospects, since social networks play a large role in many labor markets. At this stage, we do not know the implications of these biases for people's careers, but given the importance of recall in network formation, we conjecture these effects may not be small.

Data availability

The data on experiment 1 is confidential, data for experiment 2 is available upon request

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.socec.2023.102008.

References

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science, 185*(4157), 1124–1131.

Shepherd, J. W., Gibling, F., & Ellis, H. D. (1991). The Effects of distinctiveness, presentation time and delay on face recognition. *European Journal of Cognitive Psychology, 3*, 137–145.

Valentine, T. (1992). Towards an exemplar model of face processing - The effects of race and distinctiveness. *Quarterly Journal of Experimental Psychology, 44*(4), 671–703.

Kahn, S. (1993). Gender differences in academic career paths in economists. *American Economic Review, 83*(2), 52–56.

Carter, J., Fenton, S., & Modood, T. (1999). *Ethnicity and employment in higher education*. London: Policy Studies Institute.

Slone, A. E., Brigham, J. C., & Meissner, A. (2000). Social and cognitive factors affecting the own-race bias in whites. *Basic and Applied Social Psychology, 22*(2), 71–84.

Blackaby, D., & Frank, J. (2000). Ethnic and Other minority representation in UK academic economics. *Economic Journal, 110*, 293–311.

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

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

Calvó-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American Economic Review, 94*(3), 426–454.

Blackaby, D., Booth, A., & Frank, J. (2005). Outside offers and the gender pay gap: Empirical evidence from the UK academic labour market. *Economic Journal, 115* (501), 81–107.

Tibbetts, E.A. a.nd Dale, J., (2007). "Individual recognition: It is good to be different," in Tibbetts, E. A. and J. Dale 2007, *Trends in ecology and evolution*.

Fryer, R., & Jackson, M. O. (2008). A categorical model of cognition and biased decision making. *B.E. Journal of Theoretical Economics (Contributions), 8*(1). Article 6.

Belot, M. (2015). Cognitive racial discrimination: A benchmark experimental study. *Journal of Neuroscience, Psychology and Economics*.

Enke, B., Schwerter, F., & Zimmermann, F. (2020). Associative memory and belief formation (No. w26664). *National Bureau of Economic Research*.