

Geoffrey Beattie, Doron Cohen and Laura McGuire

An exploration of possible unconscious ethnic biases in higher education: The role of implicit attitudes on selection for university posts

Abstract: Despite efforts to deal with the underrepresentation of Black and Minority Ethnic (BME) staff in higher education, progress to date has been limited. We investigate the role of possible implicit attitudes towards ethnic diversity among staff and students at a leading British university. Ninety-six participants (48 White and 48 non-White) were presented with matched C.V.s of White and non-White applicants and were instructed to rate the suitability of candidates against two pre-defined job descriptions for positions at the same university (Lectureship versus Administrative role). Participants were also asked to shortlist two applicants for a subsequent interview, before completing a new multi-ethnic IAT. The new IAT assesses implicit attitudes towards BME groups as a whole, rather than focusing exclusively on a single ethnic minority. Evidence of implicit bias was observed in the IAT scores and in the White participants showing an own-race bias in terms of the proportion of Whites that they selected for the academic post, but not the administrative position. Implicit measures were a better predictor than explicit measures of actual shortlisting decisions. Policy recommendations are discussed.

Keywords: implicit-explicit attitudes; ethnic discrimination; implicit association test; shortlisting decisions; discrimination in higher education; labor market outcomes

Geoffrey Beattie: Edge Hill University. E-mail: beattieg@edgehill.ac.uk

Doron Cohen: University of Manchester. E-mail: Doroncohen@hotmail.co.uk

Laura McGuire: University of Manchester. E-mail: lauramcguire@hotmail.com

1 Introduction

Inequality in recruitment between BME groups and the majority White population continues to represent a persistent source of social and economic injustice in many Western countries (Bassanini and Saint-Martin 2008). After controlling for age, socio-economic status, and number of years in education, research shows

that BME groups, in a variety of countries, continue to face a significant “net” disadvantage in terms of gaining access to, and remaining in, the labor market (see Bassanini and Saint-Martin 2008). Evidence of labor market discrimination on the grounds of ethnicity and religion has been reported in Australia (Booth et al. 2009), Canada (Pendakur and Pendakur 1998), France (Lefranc 2010), Germany (Kogan 2011), Greece (Drydikas and Vlassis 2010), North America (Bertrand and Mullainathan 2004), New Zealand (Tobias et al. 2008), Sweden (Nordin and Rooth 2009), and the United Kingdom (Wood et al. 2009).

One technique that has been frequently used to identify labor market discrimination is “correspondence testing” (see Jowell and Prescott-Clarke 1969). In a typical correspondence test for racial discrimination, written job applications are randomly assigned either a traditional Anglo-Saxon sounding name or a BME name and are submitted for advertised vacancies. Discrimination is identified if there is a different rate of shortlisting for interview between the ethnic groups (typically White versus non-White). Bertrand and Mullainathan (2004), for instance, randomly assigned either Anglo-Saxon (e.g., Emily) or Afro-American (e.g., Lakisha) names to over 5,000 fictitious curricula vitae (hereafter C.V.s), which were then sent in response to a range of job advertisements in Boston and Chicago. They found that White candidates were 50% more likely than non-White candidates to be offered an interview. Similarly, the results of a Swedish field study (Carlsson and Rooth 2007) demonstrate that second generation Swedes are, on average, ten percentage points less likely to be invited for interview if the applicant has a Middle Eastern rather than a traditionally Swedish sounding name. In the UK research shows that BME individuals not only have to send, on average, 74% more applications than Whites to secure an interview (Wood et al. 2009), but that once in employment BME staff face lower hourly earnings and lower levels of occupational attainment, compared to equally qualified White staff (Li and Heath 2008).

1.1 Discrimination in higher education

Although research has investigated discriminatory selection practices across a broad spectrum of occupations,¹ one domain which has not been explored in any

¹ Bertrand and Mullainathan (2004): Office Management Positions, Cashier Work, Customer Services, and Administrative Support Posts. Wood et al. (2009): IT user support, IT Technician, Care Assistant, Teaching Assistant, Accountant, Accounts Clerk, Sales Assistant, Office Assistant, and Human Resource Manager.

great detail is the recruitment and retention of ethnic minority staff in Higher Education Institutes (see Leathwood et al. 2009). Nevertheless, what is clear is that despite considerable efforts to deal with the underrepresentation of BME staff among the higher academic and administrative grades of British universities, the statistics reveal that so far success has been, at best, limited (see the Equality Challenge Unit 2010). It is perhaps surprising, given the meritocratic principles on which universities are founded that the proportion of BME academic staff in UK universities is lower than that of the British working population for individuals with the appropriate qualifications for an academic post (Equality Challenge Unit 2010). This is even more surprising given that well educated individuals typically *espouse* egalitarian and liberal attitudes towards ethnic minorities (Schuman et al. 1997), at least on self-report measures.

Compared to equally qualified Whites, BME academic staff are more likely to be employed on short, fixed-term contracts in universities (34.7% for BME staff versus 28.7% for Whites), face a 2.9% median pay gap, and become increasingly underrepresented, the more senior the academic grade (Equality Challenge Unit, 2010). For example, 11% of all White academics are professors compared to just 4% of the BME academic population (Equality Challenge Unit, 2010). Indeed, according to data from the Higher Education Statistics Agency (2011) out of the 14,385 Professors in the UK, only 50 (0.35%) are Black, despite the fact that 2.8% of the population of England and Wales is Black African or Black Caribbean, according to the Office for National Statistics. In fact, the statistics reveal that 94.5% of British professors are White, 1.6% are Asian (including Asian British Indian, Asian British Pakistani, and Asian British Bangladeshi), 1.1% are Chinese, 0.7% are Other Asian, and 1.7% Other (including Black mixed ethnicity and other ethnic backgrounds), while just 0.4% are Black (Black Caribbean or Black African).

Again this appears to be an international problem with a similar picture emerging from the US, where only 5.4% of all university full-time academic staff come from a BME background (US Department of Education 2007). According to the *Journal of Blacks in Higher Education* “If we project into the future on a straight-line basis the progress of Blacks into faculty ranks over the past 26 years, we find that Blacks in faculty ranks will not reach parity with the Black percentage of the overall American workforce for another 140 years” (Leathwood et al. 2009).

1.1.1 The role of implicit and explicit attitudes

What possible biases might be at work during the consideration of candidates which could potentially explain such inequalities? One possible factor here is the

role of *implicit* attitudes, (where such attitudes are thought to be unconscious, quick, and non-reflective; see Beattie 2010), which might be impacting on a number of core processes, including initial selection interviews, promotional panels, and judgments of suitability for certain posts. In a series of articles, Greenwald and his colleagues (see Greenwald and Banaji 1995; Greenwald et al. 1998) introduced the concept of the “implicit attitude,” on the basis of significant evidence “that past experience influences judgment in a fashion not introspectively known by the actor” (Greenwald and Banaji 1995: 4). This theoretical perspective was bolstered by the introduction of the Implicit Association Test (IAT; see Greenwald et al. 1998; Nosek 2007a, 2007b). The IAT is a computer-based assessment program that measures the speed with which participants associate certain concepts (e.g., “White” and “non-White” and “good” and “bad”). The IAT does not require the participants to report their underlying attitudes and when the relationship between implicit attitudes (as revealed by the IAT) and self-report attitudes are investigated, it seems that there is often considerable divergence between the two (Beattie and Sale 2009; Beattie and Sale 2011; Beattie 2011). Explicit attitudes are those that an individual is *consciously* aware of and can articulate, whereas implicit attitudes may fall outside of an individual’s conscious control (Gregg 2008); they are thought to emerge from *automatic* and *affective* responses to a target object. Of course, self-report measures of attitudes can be skewed by a significant self-presentational bias, especially in domains that relate to socially sensitive situations, such as attitudes towards the environment (Beattie and Sale 2009; Beattie 2010), ethnicity (Park et al. 2007), and race (Nosek et al. 2002). But one important theoretical perspective holds that self-presentational bias alone cannot account for the lack of correlation and general “dissociation” between implicit and explicit attitudes.

1.1.2 The Implicit Association Test

The IAT requires the rapid categorization of various stimuli (typically words and images). In a typical race IAT respondents are asked to categorize positive and negative words and in other trials to categorise two sets of images (e.g. ‘White’ and ‘Non-White’ faces). Respondents then have to categorize individual items (e.g. an image of a ‘White’ face or a ‘Good’ word like ‘Wonderful’) into categories that have been paired (e.g. the joint category of ‘Good or Non-White’ versus the joint category of ‘Bad or White’). The more closely the two categories are associated (e.g. ‘Good’ and ‘Non-White’) the quicker and more accurate the respondents should be at assigning the individual items to the pair of categories. These pair-

ings are subsequently reversed and response latencies and error rates are calculated to produce a difference or D score (Greenwald et al. 2003).

The literature on implicit racial biases has consistently revealed that over 70% of White Americans who complete a race IAT hold some degree of implicit racial preference towards Whites over African Americans, and that this remains true even when participants claim on explicit measures to hold egalitarian and liberal values (Greenwald et al. 1998; Mitchell et al. 2003; Nosek et al. 2007). Indeed, in a large-scale internet study, Nosek et al. (2002) analyzed data collected between October 1998 and April 2000 at *Project Implicit* (www.projectimplicit.org), a purpose built website designed to provide “drop-in visitors” with the opportunity to complete a variety of IATs, including a race IAT. They reported that of the approximate 100,000 White respondents ($N = 103,316$) who completed the Race IAT, most showed a significant own race bias (Cohen’s $d = 0.83$). Indeed, White participants exhibited a *strong* preference in favor of Whites over Blacks. Furthermore, this apparent pro-White preference obtained on the implicit measure (the IAT) was considerably stronger than participants’ explicit (i.e., self-reported) preferences for Whites versus Blacks (Cohen’s $d = 0.59$). In contrast, Black respondents ($N = 17,510$) showed a strong preference for Blacks over Whites on the *explicit* measure (Cohen’s $d = -0.80$), but somewhat surprisingly Black participants exhibited a weak *pro-White* bias on the implicit measure (Cohen’s $d = 0.16$). This implicit-explicit dissociation in the responses of Black participants is thought to reflect the competing role of culturally acquired values and attitudes versus the positive associations that they feel towards their own ethnic group based on their own personal experiences (see also Jost et al. 2004). In other words, disadvantaged social groups living in a culture where they are exposed to the same racial views held by the dominant group may well acquire culturally learnt negative attitudes towards their own ethnic group, which may moderate (or even eliminate) their own race preference.

1.1.3 Critique of the Implicit Association Test

Over the past decade, the Implicit Association Test has become a widely accepted measure of implicit social cognition and has been the focus of over 450 peer-reviewed publications (Sabin et al. 2009). Part of the widespread appeal of the IAT is that it is easy to administer (especially over the internet), and it provides relatively robust findings across populations and socially sensitive domains, and the stimulus materials have been made readily available to others by Greenwald and his colleagues through *Project Implicit* (see Devine 2001). Nevertheless, there

are a number of potentially serious methodological issues with this test in terms of how it has been used. Despite the widespread popularity of the internet for recruiting large and culturally diverse samples, there is an inevitable trade-off, as Greenwald and his colleagues would readily admit (Nosek et al. 2002), between the opportunities it provides for attracting participants who might otherwise be difficult to bring into the laboratory and an unavoidable loss of experimental control. In contrast to traditional laboratory based research, online studies preclude the experimenter from controlling for the myriad of environmental cues (e.g., incoming telephone calls, ambient noise, etc.) that may lead to significant distraction among participants. Similarly, it is not possible to ascertain whether multiple data points are the product of a *single individual* (e.g., a single participant may complete the experiment multiple times), or whether a *single* data point arises from *multiple* participants (e.g., a group of friends could potentially take turns alternating between experimental trials). Furthermore, participants may intentionally (or unknowingly) provide false information when answering basic demographic questions about their own identity (e.g., age, gender, race, socio-economic status). Clearly, then, researchers must exercise caution when drawing conclusions from data collected through web-based experimental designs.

But there is another potentially more serious issue about the race IAT that impacts not just on its use over the Internet, but on its use in more controlled studies. The Race IATs presented online at *Project Implicit* fail to employ full natural-looking profile photographs (see Beattie 2013). Although the Race IAT, for instance, uses photographs of White and non-White people, the faces selected are unsmiling and appear hostile. Similarly, only a section of each face is presented which sometimes makes it challenging to determine the gender of the person featured; something that would be especially true when participants are instructed to rapidly categorize the stimuli. The photographs also appear to have been taken extremely close to the face, which itself contributes to their aggressive appearance (see Figure 1).

The Skin Tone IAT, on the other hand, uses drawings of faces that have been shaded in to change skin tone. Consequently, the images here are two-dimensional and have a somewhat synthetic appearance (see Figure 1). Part of the artificiality of these drawings is that the facial features (Euro-centric versus Afro-centric) do not change along with the skin tone, and so the faces appear very unnatural. Furthermore, there is a high degree of homogeneity in terms of the apparent age ranges of the faces used as stimuli in both the Race IAT and the Skin Tone IAT, which appear to represent exclusively young people. Using actual photographs of people of different ages, rather than drawings or partial photographs, would seem to hold greater ecological validity and be of more general applicability to everyday life. After all, it makes intuitive sense for the stimuli used in the IAT to

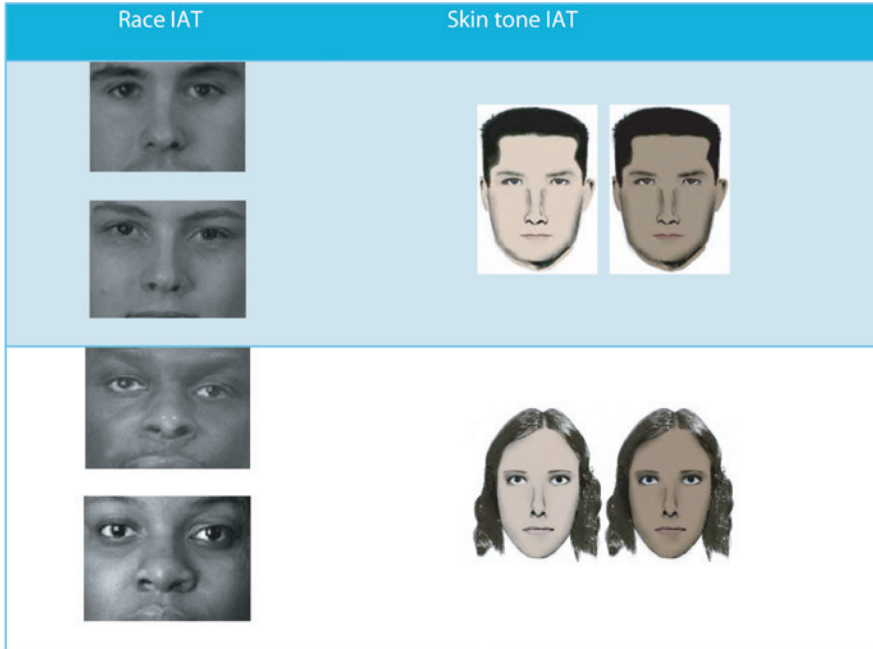


Fig. 1: Example stimuli used in the Race IAT and the Skin Tone IAT at Project Implicit

represent the kinds of faces/images that people typically encounter as they navigate their way through important decisions in everyday life.

1.2 Predictive utility of the IAT in relation to hiring decisions

Notwithstanding these limitations, a small number of studies have examined the predictive validity of the IAT in relation to employment outcomes, although none have focused on ethnic diversity within the Higher Education Sector. In a Canadian study, Son Hing et al. (2008) examined employment discrimination towards Asian job applicants. They showed that when Asian job candidates were significantly better qualified for a post than the Whites candidates, the Asians were more likely to be recommended for the position by White judges (as hopefully one would anticipate). However, when the White and Asian candidates had *similar* qualifications, the White judges were more likely to recommend White applicants. They also showed that the greater the implicit bias against Asians (as measured by the IAT), the more likely White judges were to choose the White rather than the

Asian candidate. The explicit measures, on the other hand, did not predict applicant selection.

Similarly, Rooth (2010) used an IAT to measure possible implicit biases towards job candidates with either traditional Muslim or Swedish-sounding names. Rooth analyzed the relationship between implicit and explicit attitudes and the probability that applicants would be offered an interview. Rooth found that “the probability to invite job applicants with names such as Mohammed or Ali decreases by five percentage points when the recruiter has at least a moderate negative implicit association toward Arab-Muslim men in Sweden” (2010: 529). Again, explicit measures did not predict shortlisting. The researchers go on to argue that what their results indicate is that “there are recruiters who implicitly discriminate, but who would not explicitly do so. The results present evidence that recruiting behavior is being affected by implicit prejudice rather than by explicit discrimination” (2010: 529). But, of course, this is only one specific domain. How robust is the finding, and would this generalize to other ethnic groups and different situational contexts such as that of the UK’s Higher Education Sector?

1.3 The current study

Evidently, discrimination on the grounds of race or ethnicity does represent a pervasive global and social problem. However, despite its obvious theoretical import, research attention has focused disproportionately on the *prevalence* of discrimination *per se*, rather than on the core psychological processes underpinning potential sources of bias in recruitment and hiring decisions. Research attention urgently needs to shift away from an exploration of the magnitude of “ethnic penalties” (see Heath and Li 2007), and investigate instead the psychological conditions that give rise to discrimination, and ultimately towards an understanding of *what* (if any) policy initiatives can be implemented during personnel recruitment to reduce potential bias in the decision-making processes that underpin selection. The current study is the first to examine the relationship between both implicit and explicit attitudes to ethnic minorities, and possible discriminatory selection procedures within the context of a Higher Education Institute. Rather than focusing exclusively on a *single* ethnic minority group (e.g., attitudes towards the Black community), we devised a new IAT that systematically assesses implicit and explicit attitudes towards a mixture of BME applicants (namely, Black African, Black Caribbean, Middle Eastern, Far Eastern, Japanese, Korean, Asian, and Chinese groups). The basic procedure involved asking White and non-White individuals to read through the C.V.s of four applicants for a lectureship position (lec-

turer in health psychology) and to read through a different set of four C.V.s for an administrative role within the Professional Support Services at the same university. For each post, participants were presented with two White and two non-White applicants and were asked to rate the suitability of each candidate, before selecting two for a subsequent interview. Participants were then asked to complete the new ethnic IAT. By holding the quality of each C.V. constant and by assigning a White and a non-White identity to each, this study not only enabled us to explore the role of implicit attitudes towards a broader range of job applicants, but also could reveal how discrimination might vary as a function of job type (lectureship versus administrative post).

2 Method

2.1 Participants

Ninety-six individuals (33 males, 63 females) from a leading British university, participated in the research. Half were White British, while the remainder were from various British born BME groups (i.e., Asian [43.8%], Mixed Race Asian [20.8%], Chinese [14.6%], Black African [12.5%], and Black Caribbean [8.3%]²). Most were university students.

2.2 Designing the new Ethnic IAT

In order to devise the new multi-ethnic IAT, we initially used the internet to collect a large corpus of photographs of the faces of White and non-White individuals (e.g., Black, Asian, Middle Eastern, Far Eastern, Mixed Race, etc.) representing a range of ages (18–65 years) and including both males and females. Subsequently, a number of photographs were then removed based on image quality, pose, and facial expression. Only smiling faces that were looking directly at the camera were considered. The remaining photographs were then rated independently by three judges for attractiveness on a scale from 1 to 10 where 1 corresponded to extremely unattractive and 10 corresponded to extremely attractive. From these

² These categorizations are those the participants gave themselves in response to the question “What ethnicity do you consider yourself?”

ratings only those faces that were judged to be in the middle band of attractiveness (with a mean rating of 4 to 6) were retained. The remaining photographs were then divided into two groups of 10 White and 10 non-White, ensuring that they included people from a range of ages and included an equal number of males and females. Finally the overall ratings for the photographs in the White and non-White conditions were analyzed to ensure they were comparable so that one group was not judged to be more attractive than the other.

2.3 Stimuli construction: Constructing credible C.V.s

Our stimuli were based on two job advertisements that appeared on the vacancy section of a leading British university website in August 2010. One of these vacancies advertised an academic position (*Lectureship in Health Psychology*), whilst the other advertised a role within the Professional Support Services (*Postgraduate Administrator*). The first phase of stimuli development involved constructing a series of eight credible C.V.s (four for each position). Embedded in each C.V. were a number of sections including, “level of educational attainment,” “research publication history” (lectureship post only), “previous work experience,” and “interests and pastimes” (administrative post only). Considerable attention was given to ensuring that the four C.V.s for each of the two posts were comparable in terms of the quality of the applicant (as with most C.V.s some sections were made deliberately stronger than others). For example, all of the academic applicants had studied at high ranking British universities for their first degree, but at less prestigious institutions for their second degree, they had some publications in high-impact journals, but other papers in less prestigious journals, etc. In relation to the administrative post, all candidates had a mixture of relevant and less relevant work experience with some unexplained gaps in their work history, and pastimes that focused on learning new transferable skills (e.g., attending IT classes), as well as hobbies that were less likely to appeal to a potential employer (e.g., socializing in the pub with friends). In a small number of cases typographical errors were deliberately embedded in the C.V.s in order to manipulate the perceived quality of a target C.V. (e.g., I enjoy reading the great litrature of the world).

A pre-test was conducted to ensure that the C.V.s were approximately equivalent in terms of the quality of each applicant. Ten additional participants (5 White and 5 non-White) were asked to individually rate the suitability (using a 7-point Likert scale) of each of the 8 C.V.s for the target job, on the basis of qualifications, work experience, and publication history. Critically, the C.V.s that participants

rated did not carry any information about the ethnicity of the applicant. The mean suitability ratings of C.V.s for the lectureship post was 4.74 (range 4.38 to 5.00) and 3.85 for the administrative position (range 3.67 to 4.05). These results revealed that when the C.V.s were ethnically neutral, there was little variation in how participants rated the quality of each applicant. If White and non-White participants rated the quality of the C.V.s differently when they were assigned a racial identity, then implicit biases could plausibly account for the difference.

Finally, two racial identities were randomly assigned to each of the eight C.V.s (generating 16 C.V.s in total). This was achieved by randomly assigning an Anglo-Saxon name (e.g., *Jennifer Peterson*) and an ethnic minority name to each C.V. The ethnic minority names selected were of Chinese (e.g., *Xia Lynn*), Asian (e.g., *Rezza Husseini*), and African (e.g., *Latoya Obagundu*) origin. Additionally, a passport-style photograph of a White or non-White applicant was attached to the top right hand-corner of each C.V. (see Figures 2 and 3). Applicants were matched for attractiveness and facial expressions (smiling only). A second pre-test was conducted in which an additional 20 participants (10 White and 10 non-White) were instructed to rate the attractiveness of each of the 16 faces on a 7-point Likert scale. White and non-White faces did not vary significantly by attractiveness ($T = 65$, $N = 18$, n.s.).

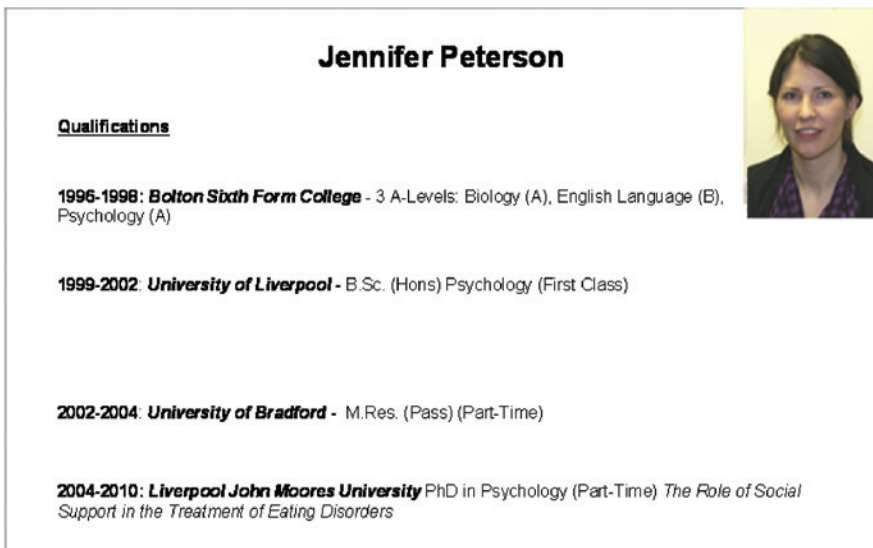


Fig. 2: Example stimuli (white lectureship post)

Ishia Nazariah

Qualifications

1996-1998: Bolton Sixth Form College - 3 A-Levels: Biology (A), English Language (B), Psychology (A)

1999-2002: University of Liverpool - B.Sc. (Hons) Psychology (First Class)

2002-2004: University of Bradford - M.Res. (Pass) (Part-Time)

2004-2010: Liverpool John Moores University PhD in Psychology (Part-Time) *The Role of Social Support in the Treatment of Eating Disorders*



Fig. 3: Example stimuli (non-white lectureship post)

2.4 Procedure

Each participant was individually presented with an electronic copy (using PowerPoint slides) of the job specification relating to one of the two positions, followed by the four C.V.s that had previously been created to fulfil the requirements of the target post. Participants were told that their task was to select the two most suitable applicants for the post. They were always shown the C.V.s of two White and two non-White applicants for each post. After reading through a C.V., participants were asked to make a judgment about the suitability of each applicant, before the next C.V. was presented. Applicant suitability was measured on a 7-point Likert scale (“1” = “not at all suitable” and “7” = “extremely suitable”). After rating the quality of all four C.V.s for a post, participants were given one minute to select two candidates that they wished to put forward for interview. They were then shown the second job description and the process was repeated. The order in which the posts were presented (academic versus administrative), the order in which the applicants appeared (White versus non-White), and the gender of the applicants were randomized throughout.

Finally, respondents completed the Ethnic IAT by following a series of computerized instructions. Two explicit (e.g., self-reported) measures of attitudes

were also taken. Respondents were asked to rate their explicit (self-reported) attitude to White versus non-White people on a 5-point computerized Likert scale, where 1 corresponds to a *strong preference for non-Whites*, 3 corresponds to *no preference*, and 5 corresponds to a *strong preference for Whites* (see Figure 4).

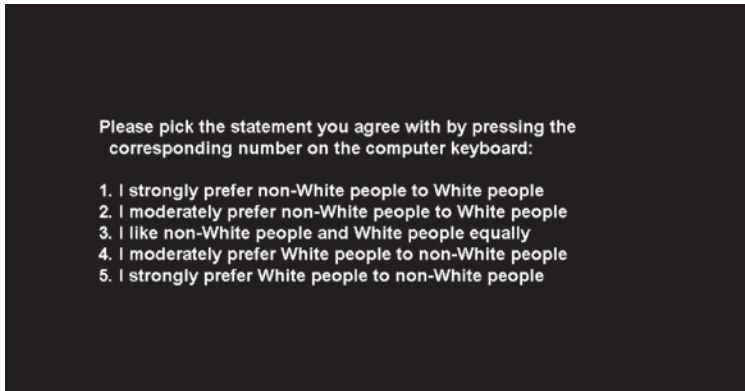


Fig. 4: A computerized version of the Likert scale for measuring explicit attitudes towards White and non-White people

Second, we assessed explicit attitudes by employing a “Feeling Thermometer,” which requires participants to rate how warm or cold they feel towards White and non-White people respectively. From the two ratings on the Feeling Thermometer, a thermometer difference (TD) score is calculated (see Figure 5). For example, a respondent with a very positive feeling towards White people might select “5” meaning “very warm” to White people and “1” meaning “very cold” to non-White people. This would yield a thermometer difference (TD) score of +4. On the other hand, a respondent who had a very positive feeling towards non-White people might select “5” meaning “very warm” towards non-White people and “1” for White people, thus producing a TD score of -4.

Finally, respondents read computerized instructions for the IAT and proceeded to complete it.



Fig. 5: A computerized version of the Feeling Thermometer for measuring explicit attitudes towards White and non-White people

3 Results

3.1 Overall implicit and explicit attitudes towards White and non-White people

The IAT was scored according to the revised scoring algorithm described by Greenwald et al. (2003), which produces a difference (D) score. In the present study, a positive D score signifies a preference for White people, and a negative D score indicates a preference for non-Whites. A D score between -0.2 and 0.2 indicates a neutral implicit attitude towards “White” and “non-White.” A ‘D’ score between 0.2 and 0.5 indicates a slight attitudinal preference towards “White,”

while scores between -0.2 and -0.5 indicates a slight attitudinal preference towards “non-White.” D scores between 0.5 and 0.8 indicate a moderate attitudinal preference towards “White,” while scores between -0.5 and -0.8 indicate a moderate attitudinal preference towards “non-White.” Finally, D scores that are greater than or equal to 0.8 indicates a strong implicit attitude towards “White,” while scores less than or equal to -0.8 indicates a strong implicit attitude towards “non-White.” These break points for “slight,” “moderate” and “strong” preferences were selected conservatively according to statistical conventions for effect size (see Cohen 1988).

Table 1 presents the overall mean explicit and implicit attitude scores.

Table 1: Mean explicit and implicit attitude scores

	Mean	SD
Likert Score	3.03	0.37
TD Score	0.33	0.75
D score	0.68	0.78

The overall mean Likert score for the new ethnic IAT was 3.03 (which indicates an almost perfect neutral explicit attitude), the mean TD score was 0.33 (which again indicates a neutral explicit attitude towards ethnicity), but critically the mean D score was 0.68 (which indicates a moderate pro-White score). In line with previous research, no significant correlation between the explicit measures and the implicit measures were observed (Likert scale, $\rho = -0.07$, $N = 96$, n.s.; TD, $\rho = -0.09$, $N = 96$, n.s.)

3.2 Effects of gender on implicit and explicit attitudes

The next analysis examined the effects of gender on explicit and implicit attitudes to White versus non-White people (see Table 2).

Table 2: Mean explicit and implicit attitude scores by gender

	Female		Male	
	Mean	SD	Mean	SD
Likert score	3.00	0.40	3.09	0.29
TD Score	0.37	0.75	0.27	0.76
D score	0.63	0.72	0.70	0.66

Table 2 demonstrates that there was very little variation in the explicit and implicit attitude scores by gender. For females, the mean Likert score was 3.00 (exactly neutral), the mean TD score was 0.37 and the mean D score was 0.63. For males, the mean Likert score was 3.09, the mean TD score was 0.27 and the mean D score was 0.70. A Mann-Whitney U test confirmed that gender had no significant effect on either measure of explicit (Likert score, $U = 952.5$, $n_1 = 63$, $n_2 = 33$, n.s.; TD score, $U = 987.00$, $n_1 = 63$, $n_2 = 33$, n.s.) or implicit attitudes (D score, $U = 1029.5$, $n_1 = 63$, $n_2 = 33$, n.s.) towards White and non-White people.

3.3 Implicit and explicit attitudes by ethnicity

The third analysis investigated whether participant ethnicity influenced implicit or explicit attitudes towards White and non-White people respectively (see Table 3).

Table 3: Mean explicit and implicit attitude scores by ethnicity

	White		non-White	
	Mean	SD	Mean	SD
Likert score	3.02	0.44	3.04	0.29
TD score	0.38	0.84	0.29	0.65
D score	0.93	0.87	0.43	0.60

Table 3 demonstrates that ethnicity had little or no effect on the explicit measures. However, the implicit attitude was affected by the ethnicity of the participant. For White participants, the mean D score was 0.93, indicating a strong pro-White preference. For non-White participants, the mean D score was 0.43, which indicates a weak pro-White preference. A Mann-Whitney U test confirmed that there was no significant difference in explicit scores by ethnicity (Likert scores, $U = 1131.5$, $n_1 = 48$, $n_2 = 48$, n.s.; TD scores, $U = 1116.5$, $n_1 = 48$, $n_2 = 48$, n.s.), but that the mean D score for White participants was significantly higher (i.e., more pro-White) than for non-White participants ($U = 660.0$, $n_1 = 48$, $n_2 = 48$, $p = 0.0001$, 2-tailed).

3.4 How did participants' ethnicity affect candidate selection for the lectureship position?

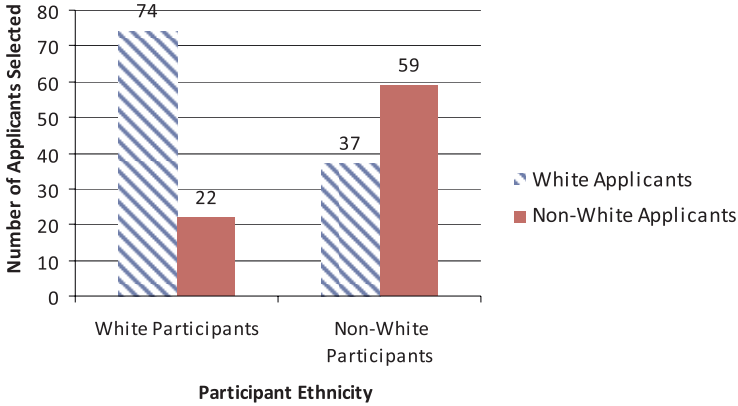


Fig. 6: Number of White and Non-White applicants selected for interview by participant ethnicity (academic post)

Figure 6 shows that despite being matched *identically* in terms of human capital White participants were around *three and a half times* more likely to select White (77.1%) than BME (22.9%) applicants for interview. On the other hand, non-White participants demonstrate a preference for BME (61.5%) candidates over White (38.5%) applicants. The next analysis examines the number of White and non-White participants as a function of the ethnicity of those candidates that they recommended for interview for the academic post (i.e., whether they recommend two White candidates, two non-Whites, or one White and one non-White applicant; see Figure 7).

Most White participants (60.4%) shortlisted two White applicants for the academic post. A third of White participants selected a White *and* a BME applicant, with very few (6.3%) recommending two BME applicants. Conversely, only around a quarter (27.1%) of non-Whites recommended two applicants from the BME population. Most non-White participants (68.8%) selected one candidate from a BME group and one from the White population, although very few (4.2%) put forward two Whites for interview.

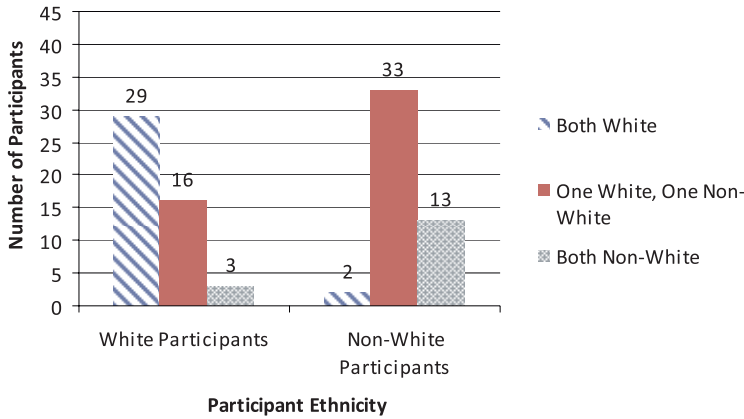


Fig. 7: Number of White and Non-White participants by recommendation to shortlist (academic post)

3.5 Relationship between IAT Score and applicant selection: Do D scores predict who will be offered an interview for the academic post?

The next stage of the analysis aimed to determine whether participants’ D scores were related to the ethnicity of the candidates selected for interview. Table 4 presents the mean D scores of participants as a function of the ethnicity of the applicants that they selected.

Table 4: Mean D scores of participants by recommendation for interview (academic post)

Participant Ethnicity	Two White	One White and one non-White	Two non-White
Whites	1.17 (SD = 0.77)	0.71 (SD = 0.86)	-0.22 (SD = 0.56)
Non-Whites	1.01 (SD = 0.00)	0.52 (SD = 0.28)	0.10 (SD = 0.99)

Irrespective of ethnicity, Table 4 demonstrates that participants who offered interviews to two White candidates had higher D scores than those who offered an interview to a White *and* a non-White applicant. Participants with the lowest D scores (i.e., the least pro-White) selected two non-White applicants. A set of three

pairwise Mann-Whitney U tests were conducted to determine whether the mean D scores varied significantly as a function of which applicants White participants recommended for interview. To reduce the risk of a type 1 error, Bonferroni correction procedures with an adjusted alpha level of 0.017 (0.05/3) were employed. White participants who shortlisted *two* White candidates were significantly more pro-White (i.e., had higher D scores) than those who selected a White *and* a non-White applicant ($U = 108.5$, $n_1 = 29$, $n_2 = 16$, $p = 0.003$, 2-tailed), or *two* non-White applicants ($U = 5.5$, $n_1 = 29$, $n_2 = 3$, $p = 0.006$, 2-tailed). Following Bonferroni corrections no significant variation in the size of the D scores was observed for Whites who selected a White *and* a non-White versus *two* non-White applicants ($U = 5$, $n_1 = 16$, $n_2 = 3$, $p = 0.033$, 2-tailed).

A second set of pairwise Mann-Whitney U tests were conducted to determine whether the mean D scores varied as a function of which applicants non-White participants shortlisted. No significant differences were observed (*Two* Whites versus one White *and* one non-White, $U = 4$, $n_1 = 2$, $n_2 = 33$, $p = 0.39$, 2-tailed; *Two* Whites versus *Two* non-Whites, $U = 2$, $n_1 = 2$, $n_2 = 13$, $p = 0.61$, 2-tailed; one White *and* one non-White, versus *Two* non-Whites $U = 125$, $n_1 = 33$, $n_2 = 13$, $p = 0.29$, 2-tailed).

3.6 Relationship between suitability ratings and applicant selection: Were the “best” candidates selected for the academic post?

Another way of examining the data is to consider the relationship between the suitability ratings and candidate selection. Did participants put forward the two applicants that they had previously judged to be the best for the post based on the 7 point rating scale of suitability? Or, were participants scoring White and non-White applicants equally, but putting forward more White (or non-White) applicants irrespective of their initial ratings? We examined this by summing the suitability ratings awarded to the two selected candidates and comparing them to the summed suitability judgments for the two unselected applicants. If participants are selecting the two candidates that they consider to be the *most suitable* for the post, then they should have awarded them significantly higher ratings, on aggregate, than unselected candidates. If however, the summed suitability ratings for the two selected candidates are not significantly higher than for the two unselected applicants, this could be taken as evidence for the existence of an underlying implicit bias.

The median rating of suitability given to White and non-White applicants by the twenty-nine White participants who put forward only White candidates was identical across the two groups (5.00 for each). A Wilcoxon Matched-Pairs Signed Ranks Test revealed that the difference in the suitability rating was not statistically significant ($T = 76.5$, $N = 20$, n.s.). In other words, these White participants were more likely to select White than non-White candidates, even though they considered all applicants to be essentially comparable in terms of their suitability for the post. A second Wilcoxon test was conducted, focusing on whether the sixteen participants who put forward a White and a non-White candidate selected those applicants that they considered to be most suitable for the post on the basis of their rating scale judgment. The median suitability of those shortlisted candidates was 5.5, while the median suitability judgment for those applicants who were not invited for interview was 4.5. The difference was statistically significant ($T = 2$, $N = 15$, $p < 0.01$, 2-tailed). In other words, this set of sixteen participants selected candidates based on some sort of merit. They put forward the two candidates that they had previously awarded the highest suitability ratings to.

In relation to non-Whites, statistical analysis revealed that those thirteen non-White participants who put forward two non-White candidates did rate the non-White applicants as more suitable than the White candidates (median rating 4.5 = Whites, non-Whites = 6.0; $T = 0$, $N = 11$, $p < 0.02$, 2-tailed). Similarly the thirty-three non-White participants who put forward both a White and a non-White applicant for interview had selected the candidates that they considered to be the most suitable for the lectureship post (median rating = 5.5 for successful applicants versus 4.5 for unsuccessful applicants; $T = 4$, $N = 27$, $p < 0.0001$, 2-tailed).

In contrast to the White participants who selected two White applicants, even though they had not rated Whites as significantly more suitable for the post than non-Whites, non-White participants were more likely to select BME candidates than Whites, because they considered them to be the “best” applicant for the post on the basis of their previously considered rating.

3.7 How did participants appraise candidates' C.V.s for the administrative position?

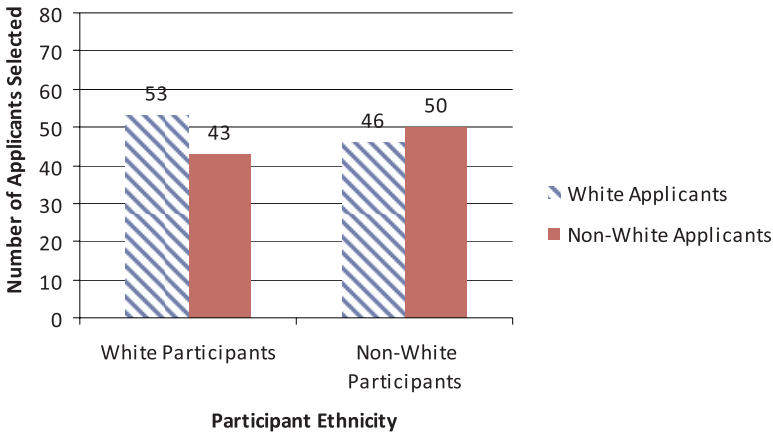


Fig. 8: Number of White and Non-White applicants selected for interview by participant ethnicity (administrative post)

Compared to the academic post, Figure 8 demonstrates that Whites and non-Whites approached parity in the number of White and BME applicants that they recommend for the administrative post. Nevertheless, there remains a race-specific bias with Whites being 10.4% more likely, on average, to be nominated for the post by Whites, and BME applicants being 4.2% more likely to be shortlisted by non-Whites. The next stage of the analysis looks at the number of White and

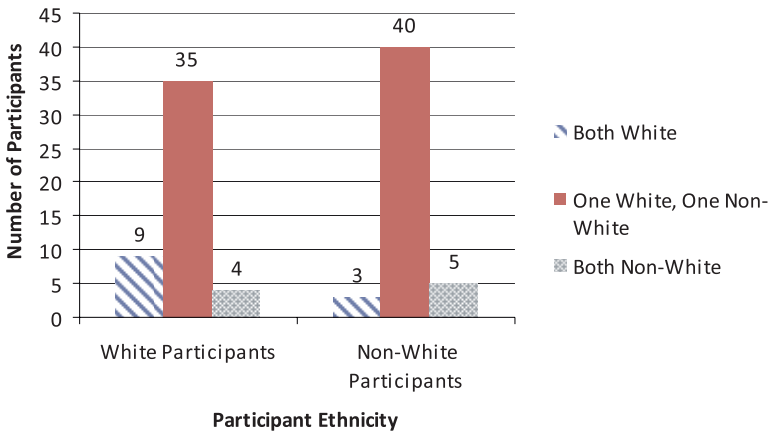


Fig. 9: Number of White and Non-White applicants selected for interview by participant ethnicity (administrative post)

non-White participants as a function of who they recommended for interview for the administrative post (see Figure 9).

Irrespective of ethnicity, Figure 9 demonstrates that participants overwhelmingly recommended a White *and* a non-White applicant for interview for the administrative post, although this pattern remains most pronounced among non-Whites (83.3 for non-Whites versus 72.9 for Whites). Almost 20.0% of White participants (18.8%) recommend *two* Whites for interview, while only half as many (10.4%) non-Whites shortlist *two* BME applicants.

3.7 Do IAT scores predict candidate selection (administrative position)?

Table 5: Mean IAT scores of White and non-White participants as a function of who they recommended for interview (administrative job)

Participant Ethnicity	Two Whites	One White, One non-White	Two non-Whites
Whites	1.82 (SD = 0.90)	0.83 (SD = 0.64)	-0.18 (SD = 0.81)
Non-Whites	0.05 (SD = 0.23)	0.53 (SD = 0.39)	-0.13 (SD = 0.56)

In line with predictions, White participants who put forward *two* Whites were significantly more pro-White than Whites who recommended a White *and* a non-White applicant ($U = 39$, $n_1 = 9$, $n_2 = 35$, $p = 0.01$, 2-tailed). However, after Bonferroni corrections, no significant differences in the size of the D scores were observed for Whites who selected a White *and* a non-White versus *two* non-White candidates ($U = 0$, $n_1 = 9$, $n_2 = 4$, $p = 0.03$, 2-tailed), or for those who selected *two* Whites versus *two* non-Whites ($U = 25$, $n_1 = 35$, $n_2 = 4$, $p = 0.04$, 2-tailed). Conversely, non-White participants who recommended one candidate from a White *and* one from a BME background were significantly more pro-White than those who shortlisted *two* non-Whites ($U = 27.5$, $n_1 = 40$, $n_2 = 5$, $p = 0.006$, 2-tailed). However, after Bonferroni corrections, no significant differences were observed in the size of the D score between non-Whites who recommended *two* Whites versus *two* non-Whites ($U = 6$, $n_1 = 3$, $n_2 = 5$, $p = 0.79$, 2-tailed), or *two* Whites versus one White *and* one non-White ($U = 16.5$, $n_1 = 3$, $n_2 = 40$, $p = 0.03$, 2-tailed).

3.8 Relationship between suitability ratings and applicant selection: Were the “best” candidates selected for the administrative post?

Statistical analyses revealed that the thirty-five participants who put forward one White and one non-White candidate were significantly more likely to shortlist those applicants that they considered to be most suitable for the post (median suitability for selected candidates was 5.5 versus 4.0 for unselected candidates; $T = 64$, $N = 26$, $p < 0.01$, 2-tailed). The suitability ratings of White and non-White applicants by the nine White participants who put forward *two* White candidates were examined. The median rating of suitability was 5.5 for the White applicants and 4.5 for the non-Whites. No significant differences were observed ($T = 10.0$, $N = 9$, n.s.). Given that prospective applicants differ only in terms of name and face, this again suggests that some sort of ethnic bias is a major factor in why these nine White participants recommend two White applicants, even though they had not rated them as significantly more suitable than the two non-Whites for the post. Of course, one problem with this conclusion is that here the sample size is small.

Finally, statistical analysis revealed that the 40 non-White participants who put forward both a White and a non-White applicant selected the candidates that they considered to be significantly more suitable for the administrative post (the median suitability rating for the successful applicants = 5.5 versus 4.5 for the unsuccessful applicants) ($T = 136$, $N = 30$, $p = 0.05$, 2-tailed).

4 Discussion

This study aimed to fill an important lacuna in the psychological literature by exploring, for the first time, the impact of ethnic attitudes on personnel selection among staff and students at a leading British University. Consistent with previous research on implicit racial attitudes (Greenwald et al., 1998; Mitchell et al. 2003; Nosek et al. 2007), we observed that irrespective of ethnicity, participants across the sample held a moderate pro-White bias ($D = 0.68$). In line with the wider literature, considerable divergence between implicit and explicit attitudes was also observed (see Nosek et al. 2002; Park et al. 2007; Beattie and Sale 2009; Beattie 2010, 2011, 2013; Beattie and Sale 2011). Despite the fact that explicit scores from both the Likert scale ($M = 3.03$) and the Feeling Thermometer ($M = 0.33$) clearly indicated that participants had no explicit preference for White over non-White people, measures of implicit attitudes (as revealed by the IAT) showed that White

participants exhibited a strong pro-White bias (D score = 0.93), while non-White participants demonstrated a weak pro-White preference (D score = 0.43).

Although there was a demonstrable ethnic bias in relation to *which* candidates were shortlisted for each of the two posts, with both groups showing an own-race preference, our results provide clear evidence that implicit, (although not explicit), ethnic attitudes predicted the shortlisting decisions of *White*, but not *non-White* participants. Indeed, in terms of their selection decisions, White participants were *ten times* more likely to shortlist two White candidates for the *lectureship* post than two non-White candidates, despite the fact that the C.V.s were matched identically in terms of human capital. However, they typically selected a White and a non-White applicant for (what most would consider to be) the less prestigious *administrative* post. In contrast, non-White participants tended to shortlist one White and one non-White candidate for both the *lectureship* and the *administrative* posts. Given that applicants' C.V.s varied solely by name and face, it seems that the only plausible explanation for the disproportionate number of White applicants selected by White participants is an underlying ethnic bias.

What is striking, however, is that the candidate suitability ratings awarded by White participants who selected two White candidates for the *lectureship* post were comparable (i.e., did not vary significantly) across White and non-White applicants. In other words, when participants are provided with a relatively *constrained* task in which they have to systematically analyze each applicant's C.V., by comparing their qualifications and competencies against a pre-defined set of job specifications, they typically arrive at a more rational and accurate conclusion about the applicant's suitability for the post. That is, they do not rate White and non-White applicants differently, as one hopefully might expect given that the C.V.s were matched across the corpus. It is only when they are subsequently placed under *explicit* time-pressure (i.e., given 60 seconds) and instructed to perform a relatively *open-ended* task (i.e., shortlist two applicants for interview) that a significant proportion of White participants disregarded their previous suitability judgments and shortlisted two applicants from their own ethnic background.

These findings are consistent with previous research, which shows that when participants, in a number of domains, make decisions under time-pressure, implicit biases are a better predictor of behavioral choice than explicit measures (Fries et al. 2006; Beattie and Sale 2011). The data provide compelling support for the idea that time pressure moderates the predictive utility of implicit versus explicit attitudes, presumably because it precludes the possibility of deliberative and extensive information processing. Consequently, time-pressured decisions tend to be guided by highly accessible "implicit" preferences (see also Fazio and

Towles-Schwen 1999), which increases people's vulnerability to the automatic activation of cultural stereotypes.

This is a stark finding for at least two reasons. First, research demonstrates that employers often read through C.V.s while they are under time-constraints, devoting an average of anywhere between thirty seconds and two minutes to each application (McGee 2003), although the process is likely to take longer when analyzing the C.V.s of academic staff. Second, due to the enduring underrepresentation of BME staff in senior academic positions, selection panels are likely to consist of predominantly of White individuals who, as the current data reveal, might well hold strong pro-White preferences, despite being well educated, working in a culturally diverse institution and publically espousing egalitarian and liberal values. Taken collectively, these results may go some way in accounting for the underrepresentation of BME staff in the Higher Education Sector.

Overall, IAT scores were useful predictors of behavioral choice. The larger the mean D scores of White participants, the more likely they were to disregard their previous ratings of suitability and put forward two White or a White *and* a non-White applicant, even though they had assessed the quality of both candidates to be comparable. However, the D scores of non-Whites did not predict who they nominated for subsequent interviews. One explanation for this asymmetry is that because the non-White group were essentially heterogeneous, comprising multiple ethnicities (e.g., Asian, Black, and Chinese applicants), there is no plausible theoretical reason to assume that BME participants would have a non-White bias for anything other than their *own* specific ethnic sub-group. Moreover, if non-White people really do have a weak underlying pro-White bias, as both Greenwald's data and the data presented here suggest, then it would seem that their implicit biases around race or ethnicity are closer to neutral than for White participants, allowing non-Whites to consider White and non-White applicants more fairly. After all, the majority of non-Whites appeared to select candidates for both positions on merit. They typically put forward the two candidates that they had previously rated as most suitable for the post.

4.1 Implications for policy

The research presented here suggests a number of targeted interventions that may be useful in reducing the likelihood (and magnitude) of possible bias occurring during the first stage of labor market discrimination (the consideration of an applicant's C.V.). One possible starting point here might be to disrupt the "natural connection" between implicit process and behavior. Indeed, the results suggest

that where possible, C.V.s should be made anonymous by removing identifying information such as an applicant's name and personal details prior to being considered by a selection panel. Of course, there are many domains in which blanket anonymity would be impractical. Taking academia as a case in point, it would be almost impossible for applicants to hide their identity from a selection panel, as prospective employers are likely to want to review the number and quality of a candidate's publications. Nevertheless, for many other posts (including administrative roles), this measure could be effectively implemented, as it is in many countries world-wide (see also Suk 2007). Similarly, another policy recommendation that could be implemented during recruitment might involve each member of the selection panel independently rating the suitability of a candidate on a 7-point scale, using a pre-defined set of job specific criteria. The summed totals of the panels' ratings would then serve as the guide to job selection, rather than selecting candidates on something as vague or open ended as "gut feelings" (see Gladwell 2005). Finally, assembling an ethnically diverse selection panel (although this might prove challenging due to the underrepresentation of non-Whites) would reduce bias, because the data suggests that these individuals hold implicit attitudes closer to neutral than most White participants.

Another approach to reducing bias during C.V appraisal involves changing the social organization of selection panels and interviews. The more time pressure, the more complex the decision making is in terms of what is expected of a selection panel, the more tasks the panel are being given simultaneously, the more "intensely" they are chaired with *shorter latencies* of response between turns and more overlapping talk by the panel members (see Beattie 1983), the less time there will be for individual reflection and decision-making and the more likely the outcomes will be biased in particular ways. Consequently, panel members must be given sufficient time to reflect on the relative merits (and weaknesses) of each candidate. Selection panels must be more "relaxed" affairs, at least in terms of temporal patterning (see Beattie 1983).

Beyond this, research must actively investigate how to prevent implicit biases from perpetuating labor market inequalities. Kandola (2009) has argued that so-called "*implementation intentions*," (a form of contingent self-talk) might offer one possible solution to tackling these complex issues. He claims that implementation intentions specify:

where, when and how we will behave in a particular way. It's not exactly programming, and it's certainly not brain-washing. But it is a way of directing our attention to particular contexts in which we are likely to act from unconscious habit rather than conscious intention. We train ourselves to recognize certain situations in which we want to act differently. (Kandola 2009: 175)

According to Kandola, the benefit of these implementation intentions is that they allow “a kind of memory of the future. The pattern laid down by the plan creates a strong association between environmental cues and intended behavior . . . When the anticipated situation arises, the memory traces formed by the plan are readily accessed and the desired behavior is activated” (2009: 175). In other words, what Kandola appears to be arguing is that the “natural” connection between a situation and the automatic behavior that arises because of implicit processes can be disrupted by using a degree of pre-planning. Kandola tried to use such implementation intentions to remove unconscious race bias. One of the groups in his study was an implementation intention group and they were given an implementation intention as well as a goal. The goal that they were given was “Don’t be prejudiced” and the implementation plan was “If I see a dark face then I’ll ignore skin color.” According to Kandola “those participants given an implementation plan performed significantly different to those in the control group. Their IAT scores were reduced to near zero meaning an absence of prejudice” (2009: 177). One problem with Kandola’s claim is that it seems never to have been published (it was merely an oral presentation at a conference). Consequently, exactly how effective they are for preventing implicit bias does need to be carefully evaluated.

Nevertheless, following Kandola’s logic, a number of implementation intentions could potentially be employed. Panel members, for instance, might be asked to reflect on the following implementation intervention: “*If I notice that I have only shortlisted applicants from a particular ethnic or racial group, when the applicants come from a range of ethnic and racial backgrounds, then I will look at the C.V.s once again to make sure that some sort of implicit process were not at work in my decision making*” – prior to shortlisting candidates. Similarly, raising awareness of implicit attitudes among academics and employers more generally might be sufficient to mitigate some of these biases.

But perhaps the greatest challenge of all requires researchers to develop even more focused interventions that attempt to change an individual’s underlying *implicit* attitude. Olson and Fazio (2006) have suggested using Implicit Evaluative Conditioning to reduce automatically activated individual prejudice. Implicit Evaluative Conditioning is a non-conscious learning process where, for example, experimental participants believe they are taking part in an experiment about “attention and surveillance,” in which they are presented with a stream of “randomly assembled images” on a computer screen. In reality, on the computer screen various images are paired in particular ways such that images of Black or White individuals appear on the screen in association with either positive or negative words or images. The basic idea is that the pairing the stimuli produces a certain response in participants and through repeated association this process

starts to produce a similar effect on its own, and this all happens below the level of conscious awareness. Olson and Fazio claim that this type of associative learning does reduce automatically activated racial attitudes and that they stay low for at least two days following the experiment. However, little is known about the long-term effectiveness of this sort of attitude-change intervention.

4.2 Conclusions

The present research provides a starting point to better understand the complex interplay between implicit and explicit processes, and their potential role in recruitment, within Higher Education Institutes. It does, of course, leave many questions unanswered and future research must begin to explore the cognitive processes, including the implicit cognitive processes, which take place during real selection processes themselves, as opposed to mere experimental surrogates. One potentially important aspect of such processes is what selectors 'see' when they are faced with candidates from different ethnic backgrounds, including what they see when they look at their C.V.s. Indeed, precisely how employers' visual attention might be guided by their implicit biases when they read through different candidates' C.V.s remains something of an open question that warrants immediate and urgent research attention (but see Beattie 2013 for, at least, a start in this direction). After all there is typically a lot of different sorts of information on C.V.s that can be focussed on and fixation on different aspects of a C.V. could result in very different 'rational' conclusions about the merits of any given candidate.

But what is abundantly clear is that many challenges need to be overcome in order to tackle labor market inequalities related to ethnicity. The research presented here has made the first tentative steps towards investigating some of these complex but crucially important issues.

References

- Bassanini, Andrea & Anne Saint-Martin. 2008. The price of prejudice: Labor market discrimination on the grounds of gender and ethnicity. In *OECD employment outlook*. Paris: OECD Publishing.
- Beattie, Geoffrey. 1983. *Talk: An analysis of speech and non-verbal behavior*. Milton Keynes: Open University Press.
- Beattie, Geoffrey. 2010. *Why aren't we saving the planet? A psychologist's perspective*. London & New York: Routledge.
- Beattie, Geoffrey. 2011. *Get the edge: How simple changes will transform your life*. London: Headline.

- Beattie, Geoffrey. 2013. *Our racist heart? An exploration of unconscious prejudice in everyday life*. London: Routledge.
- Beattie, Geoffrey & Laura Sale. 2009. Explicit and implicit attitudes to low and high carbon footprint products. *International Journal of Environmental, Cultural, Economic and Social Sustainability* 5. 191–206.
- Beattie, Geoffrey & Laura Sale. 2011. Shopping to save the planet? Implicit rather than explicit attitudes predict low carbon footprint consumer choice. *International Journal of Environmental, Cultural, Economic and Social Sustainability* 7. 211–232.
- Bertrand, Marianne & Sendil Mullainathan. 2004. Are Emily and Gregg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94. 991–1013.
- Booth, A., A. Leigh & E. Vaganova. 2009. *Does racial and ethnic discrimination vary across minority groups?* Unpublished paper.
- Carlsson, Magnus & Dan-Olaf Rooth. 2007. Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labor Economics* 14. 716–729.
- Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*, 2nd edn. Hillsdale, NJ: Lawrence Erlbaum.
- Devine, Patricia. 2001. Implicit prejudice and stereotyping: How automatic are they? Introduction to the special section. *Journal of Personality and Social Psychology* 81. 757–759.
- Devine, Patricia. G. 2005. Prejudice with and without compunction: Allport's inner conflict revisited. In J. F. Dovidio, P. Glick & L. A. Rudman (eds.), *On the nature of prejudice: Fifty years after Allport*, 327–342. Oxford: Blackwells.
- Drydakis, Nick & Minas Vlassis. 2010. Ethnic discrimination in the Greek labor market: Occupational access, insurance coverage and wage offers. *Manchester School* 78: 201–218.
- Equality Challenge Unit (2010). *Equality in higher education: statistical report 2010*. <http://www.ecu.ac.uk/publications/equality-in-he-stats-10> (accessed 27 June 2013).
- Fazio, Russell H. & Tamara Towles-Schwen. 1999. The MODE model of attitude-behavior processes. In S. Chaiken & Y. Trope (eds.), *Dual process theories in social psychology*, 97–116. New York: Guilford.
- Friese, Malte W., Michaela Wänke & Henning Plessner. 2006. Implicit consumer preferences and their influence on product choice. *Psychology and Marketing* 23. 727–740.
- Gladwell, Malcolm. 2005. *Blink: The power of thinking without thinking*. New York: Little Brown.
- Govan, Cassie L. & Kipling D. Williams. 2004. Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal of Experimental Social Psychology* 40(3). 357–365.
- Greenwald, Anthony G. & Mahzarin R. Banaji. 1995. Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review* 102. 4–27.
- Greenwald, Anthony, Debbie McGhee & Jordon Schwarz. 1998. Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology* 74. 1464–1480.
- Greenwald, Anthony G., Brian A. Nosek & Mahzarin R. Banaji. 2003. Understanding and using the Implicit Association Test: 1. An improved scoring algorithm. *Journal of Personality and Social Psychology* 85. 197–216.
- Gregg, Aiden P. 2008. Oracle of the unconscious or deceiver of the unwitting? *Psychologist* 21. 762–766.
- Heath, Anthony & Yaojun Li. 2007. Measuring the size of the employer contribution to the ethnic minority employment gap. *Consultation paper for the National Employment Panel*. London.

- Jost, John T., Mahzarin R. Banaji & Brian A. Nosek. 2004. A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering the status quo. *Political Psychology* 6. 881–919.
- Jowell, Roger & Patricia Prescott-Clarke. 1970. Racial discrimination and White-collar workers in Britain. *Race and Class* 11. 397–417.
- Kandola, Binna. 2009. *The value of difference: Eliminating bias in the organization*. Oxford: Pearn Kandola.
- Kogan, Irena. 2011. New immigrants – Old disadvantage patterns? Recent immigrants into Germany. *International Migration* 49. 91–117.
- Leathwood, Carol, Uvanney Maylor & Marie-Pierre Moreau. 2009. *The experience of black and minority ethnic staff working in higher education*. London: Equality Challenge Unit.
- LeFranc, Arnaud. 2010. Unequal opportunities and ethnic origin: The labor market outcomes of second-generation immigrants in France. *American Behavioral Scientist* 53. 1851–1882.
- Li, Yaojun & Anthony Heath. 2008. Minority ethnic men in British labour market (1972–2005). *International Journal of Sociology and Social Policy* 28. 231–244.
- McGee, Paul. 2003. *Writing a C.V. that works: How to develop and use your key marketing tool*. Cromwell Press: Wiltshire.
- Mitchell, Jason P., Brian N. Nosek & Mahzarin R. Banaji. 2003. Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General* 132. 455–469.
- Nordin, Martin & Dan-Olof Rooth. 2009. The ethnic employment and income gap in Sweden: Is skill or labor market discrimination the explanation? *Scandinavian Journal of Economics* 111. 487–510.
- Nosek, Brian A. 2007a. Implicit-explicit relations. *Current Directions in Psychological Science* 16. 65–69.
- Nosek, Brian A. 2007b. Understanding the individual implicitly and explicitly. *International Journal of Psychology* 42. 184–188.
- Nosek, Brian A., Mahzarin R. Banaji & Anthony G. Greenwald. 2002. Harvesting implicit group attitudes and beliefs from the demonstration website. *Group Dynamics* 6. 101–115.
- Nosek, Brian A., Frederik L. Smyth & Jeffrey J. Hansen. 2007. Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology* 1. 1–53.
- Olson, Michael A. & Russel H. Fazio. 2006. Automatically-activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin* 32. 421–433.
- Park, Jaihyun, Karla Felix & Grace Lee. 2007. Implicit attitudes toward Arab-Muslims and the moderating effects of social information. *Basic and Applied Social Psychology* 29. 35–45.
- Pendakur, Kisha & Ravi Pendakur. 1998. The color of money: Earnings differentials among ethnic groups in Canada. *Canadian Journal of Economics* 31. 518–548.
- Rooth, Dann-Olof. 2010. Automatic associations and discriminations in hiring: Real world evidence. *Labor Economics* 17. 523–534.
- Sabin, Janice A., Brian A. Nosek, Anthony G. Greenwald & Frederick P. Rivara. 2009. Physicians' implicit and attitude about race by MD race, ethnicity, and gender. *Journal of Health Care for the Poor and Underserved* 20. 896–913.
- Schuman, Howard, Charlotte Steeh & Lawrence Bobo. 1997. *Racial attitudes in America: Trends and interpretations*. Cambridge, MA: Harvard University Press
- Son Hing, Leanne S., Greg A. Chun-Yan, Leah K. Hamilton & Mark P. Zanna. 2008. A two-dimensional model that employs explicit and implicit attitudes to characterize prejudice. *Journal of Personality and Social Psychology* 94. 971–987.

- Suk, Julie C. 2007. Discrimination at will: Job security protections and equal employment opportunity in conflict and equal employment opportunities in conflict. *Stanford Law Review* 60. 73–114.
- Tobias, Martin, Aloka Bhattacharya, & Paul White. 2008. Cross classification of the New Zealand population by ethnicity and deprivation: Trends from 1996 to 2006. *Australian and New Zealand Journal of Public Health* 32. 431–436.
- Wood, Martin, Jon Hales, Susan Purdon, Tanja Sejersen & Oliver Hayllar. 2009. A test for racial discrimination in recruitment practice in British cities. *Department of Work and Pensions* (Research Report No 607).

Bionotes

Geoffrey Beattie is Professor at Edge Hill University and a Masters Supervisor at the University of Cambridge (beattieg@edgehill.ac.uk). His research interests include the relationship between gesture and speech, implicit attitudes towards the environment and the role of implicit processes connected to ethnicity in everyday social life. His publications include “Possible unconscious bias in recruitment and the need to promote equality” (with P. Johnson, 2011); “Making an action film. Do films such as Al Gore’s *An Inconvenient Truth* really make any difference to how we think and feel about climate change?” (2011); and *Our racist heart? An exploration of unconscious prejudice in everyday life* (2012).

Doron Cohen is a research associate at the University of Manchester (Doroncohen@hotmail.co.uk). His research interests include the relationship between speech and gesture and the impact of implicit attitudes on behavior. His publications include “Nonverbal indicators of deception: how iconic gestures reveal thoughts that cannot be suppressed” (with G. Beattie & H. Shovelton, 2010); and “Tracking the distribution of individual semantic features in gesture across spoken discourse: New perspectives in multi-modal interaction” (with G. Beattie & H. Shovelton, 2011).

Laura McGuire was a research assistant at the University of Manchester (lauramcguire@hotmail.com). Her research interests include the impact of implicit attitudes on unconscious eye movements, and emotional responses to messages in the context of sustainability. Her publications include “Do we actually look at the carbon footprint of a product in the initial few seconds? An experimental analysis of unconscious eye movements”; “An Inconvenient Truth? Can extracts of film really affect our psychological mood and our motivation to act against climate change?”; “Are we too optimistic to bother saving the planet? The relationship between optimism, eye gaze and negative images of climate change” and “See no evil? Only implicit attitudes predict unconscious eye movements towards images of climate change”.

