Oxford Review of Economic Policy, 2024, **40**, 579–591
https://doi.org/10.1093/oxrep/grae032
Article

Do ethnic minorities have incentives to omit photographs from résumés? Experimental evidence from Germany and the Netherlands

Mariña Fernández-Reino* and **Martí Rovira****

*Centre on Migration, Policy and Society, University of Oxford, UK, e-mail: marina.fernandez-reino@compas.ox.ac.uk
**Department of Political and Social Sciences, Universitat Pompeu Fabra, Barcelona, Spain, e-mail: marti.rovira@upf.edu

The authors declare that financial support was received for the research, authorship, and/or publication of this article. Funded by European Union (Equalstrength, Horizon Europe-CLS-2022-Transformations-01 call, project number 101094527); UKRI Innovate UK (Horizon Europe Guarantee, project number 10066411); European Union-NextGenerationEU, Ministry of Universities and Recovery, Transformation and Resilience Plan, through a call from Pompeu Fabra University (Barcelona).

Abstract

Ethnic minority job-seekers may have incentives to include or omit photographs from résumés depending on the social norms regarding this practice and the extent to which their outward appearances diverge from the majority group. To investigate minorities' incentives to engage in a specific type of résumé whitening—omitting photographs to avoid discrimination based on phenotypical traits or religious attire—we draw on data from an audit field experiment conducted between 2016 and 2018 in the Netherlands and Germany. Our analysis highlights the challenges for anticipating discrimination (and, thus, deciding to engage in résumé whitening) since it is difficult to predict employers' true preferences with respect to the inclusion of photographs. In the Netherlands, unveiled Turkish women were better off submitting their picture rather than omitting it, but the premium associated with the inclusion of photographs disappeared if they wore the Muslim headscarf. For unveiled Moroccan women, there was no discernible advantage or disadvantage associated with the inclusion of pictures in job applications. In contrast, in the German context, minority women with high prototypical appearance (brown phenotype or wearing the Muslim headscarf) would have incentives to omit their photographs from job applications, particularly if they wore the Muslim headscarf. For Moroccan and Turkish men in the two countries, the inclusion of pictures with different phenotypes does not affect employers' responses. To conclude, we explore the advantages and disadvantages of blind recruitment policies (i.e. omitting all ethnic identifying cues from résumés, including photographs and names) in the European context.

Keywords: field experiment, audit study, ethnic discrimination.

JEL codes: C93, J7, Z13

I. Introduction

In 1966, the UK government commissioned a study on racial discrimination in England led by W.W. Daniel (1968). The research team initially planned to survey minorities about the discrimination they experienced. In the end, they decided to implement a covert 'audit study' with actors applying to real-world jobs, as they were concerned that migrants' reports would not convince sceptics of the existence of widespread discrimination. The results revealed stark racial disparities in callback rates: only 2.5 per cent of Black or Asian testers received a positive response from employers compared to 37.5 per cent of White testers. These findings played a crucial role in the passage of the 1968 Race Relations Act in the UK, which marked the first legal prohibition of employment and housing discrimination in the country (Rich, 2018; Heath and Di Stasio, 2019; Quillian and Midtbøen, 2021, p. 392). As a result of the early adoption of this anti-discrimination legislation, the practice of attaching pictures to job applications gradually fell out of favour and is currently perceived as unprofessional in the British context. Additionally,

requesting job candidates to attach a photograph can breach anti-discrimination laws (Equality and Human Rights Commission, 2011, p. 229).

In contrast, in central and southern European countries, the social norms regarding the inclusion of photographs in job applications differ from those in the UK. For instance, in Germany, it is a standard practice to include a professional-looking photo on résumés even though employers are not allowed to request photographs (Kanning, 2016; Weichselbaumer, 2017). The distinct normative expectations associated with the inclusion of photographs in job applications not only mirror diverse approaches to the recruitment process across European countries, but also reflect historical differences in the implementation of anti-discrimination policies and the salience of social norms to prevent racism and ethnic discrimination (Simon, 2017).

In contexts where the inclusion of photographs in job applications is an accepted practice, non-White ethnic minority applicants may choose to omit them if they think that making their ethnicity or race immediately visible to employers could lower their prospects of being hired. Indeed, a headshot is undeniably the most conspicuous signal of a person's race or ethnicity that could be included in a résumé, as individuals rely on phenotypic facial traits for making ethnic and racial categorizations (Maddox, 2004; Skinner and Nicolas, 2015). Moreover, the inclusion of photographs can facilitate discrimination in the early stages of the hiring process based on other visible traits such as religious affiliation (e.g. bias against applicants wearing the Muslim headscarf), attractiveness, or weight (Rooth, 2009; Ruffle and Shtudiner, 2015; Rich, 2018; Fernández-Reino *et al*., 2023). Research on colourism and race prototypicality bias has shown that lighter skin tones and Eurocentric facial features are more socially valued in Western societies than darker skin tones and non-White phenotypes (Maddox, 2004; Marira and Mitra, 2013; Dixon and Telles, 2017), which creates an incentive for minority individuals with non-European phenotypic traits to omit their photographs from their résumés. Indeed, there is evidence that ethnic and racial minorities engage in 'résumé whitening'—concealing or downplaying the salience of ethnic and racial cues (Kang *et al*., 2016)—to boost the probability of being hired (Ruedin and Van Belle, 2023).

In this article, we use publicly available data from the first cross-national harmonized field experiment on hiring discrimination—the GEMM study (Lancee *et al*., 2021)—to assess the impact of omitting photographs from job applications on the probability of ethnic minorities receiving positive callbacks from employers. Specifically, we investigate the degree to which minorities would have incentives to omit or include headshot pictures in their résumés depending on their outward appearance, i.e. phenotypic traits and religious attire. We focus on two European countries, Germany and the Netherlands, in which the practice of including profile pictures on résumés is common (Weichselbaumer, 2017; Aureus, 2024).

Recently published articles based on data from the GEMM study examined employers' phenotypic discrimination (see Polavieja *et al*. (2023) on the impact of race and ethnicity on employers' behaviour) and anti-Muslim discrimination (see Fernández-Reino *et al*. (2023) on the penalty against women wearing the Muslim headscarf). In this article, however, our primary objective is not to measure employers' biases against applicants with various phenotypes and minority women who wear the *hijab*. Rather, our aim is to investigate whether, based on the data, there are incentives for ethnic minorities to engage in a specific type of résumé whitening—submitting a résumé without a headshot photograph—depending on their phenotypic traits and whether they wear the Muslim headscarf. We conclude by discussing the implications of our findings for anti-discrimination interventions in the hiring process and the limitations of the different approaches.

II. The need for field experiments to measure employers' behaviour

Departing from Daniel's seminal report (1968), 'audit studies', also known as 'field experiments', have become the gold standard for measuring ethnic and racial discrimination in the labour market. In their current format, online applications of equally qualified candidates of different ethnicities or races (signalled by forenames and surnames and, depending on the national context, headshot pictures too) are submitted to real job vacancies. The disparity in callback rates between majority and minority equally qualified candidates is then used as a reliable metric of discrimination that is not affected by social desirability bias, i.e. employers hiding their preferences or prejudice in surveys or interviews (Bertrand and Duflo, 2017; Gaddis, 2019; Quillian and Midtbøen, 2021). Field experiments are not without limitations, primarily concerning their ability to detect discrimination only in the initial stages of the selection process and its limited external validity, i.e. most studies are only able to measure discrimination in a limited number of occupations and job vacancies, which limits the generalizability of the findings to the entire labour market. Despite these limitations, field experiments remain superior to other approaches in providing a reliable measure of employers' direct discrimination against minority applicants (see Pager, 2007; Quillian and Midtbøen, 2021).

To date, over 140 field experiments have been conducted across 30 countries (Rich, 2014; Quillian and Midtbøen, 2021; Lippens *et al*., 2023). In the US, most of this research has concentrated on discrimination against African Americans and Hispanics (e.g. Bertrand and Mullainathan, 2004; Pager *et al*., 2009; Pager and Karafin, 2009). In Europe, correspondence audits conducted in the labour market have primarily compared employers' behaviour towards native majorities and the largest minorities of immigrant descent, revealing pervasive discrimination against Black and Muslim minorities (Zschirnt and Ruedin, 2016; Heath and Di Stasio, 2019; Quillian *et al*., 2019; Di Stasio *et al*., 2021; Quillian and Midtbøen, 2021; Thijssen *et al*., 2022; Fernández-Reino *et al*., 2023; Lippens *et al*., 2023; Quillian and Lee, 2023; Rovira, 2023).

III. Do ethnic minority applicants have incentives to omit their photographs from their résumés?

Ethnic minority applicants, especially those from the most stigmatized groups, have incentives to 'whiten' their résumés due to the prevalence of ethnic discrimination in labour markets (Ruedin and Van Belle, 2023). Résumé whitening involves concealing or downplaying racial or ethnic cues, such as avoiding ethnic-sounding names or excluding photographs. Although some minority individuals are critical of these practices, a field experiment conducted in the US labour market showed that whitening résumés substantially increased the callback rates of minority applicants (see Kang *et al*., 2016).

The incentives to engage in a specific type of résumé whitening—omitting headshot pictures from résumés—are probably stronger for ethnic minorities who are both negatively stereotyped and visible due to their phenotypic traits or religious attire. Prior research showed that phenotypic traits (e.g. skin tone) are more relevant than other ethnic cues (e.g. names) for making ethnic or racial categorizations and prejudicial attitudes (Skinner and Nicolas, 2015). Outward appearances, including phenotypic features and attire, trigger both cognitive processing and affective responses. First, outward appearances serve to classify others into distinct (race, ethnic or religious) categories. In this regard, it has been shown that people have mental representations or prototypes of (race, ethnic or religious) categories which are used to evaluate others based on the extent to which they mirror the stereotypical image of category members (Maddox, 2004). Thus, minorities with non-stereotypical traits (i.e. with low prototypicality or low prototypical features) are less likely to be identified as such. Second, outward appearances have affective significance for people, with certain physical traits triggering positive or negative reactions (Livingston and Brewer, 2002). For example, minorities with features that are considered prototypical of their minority group (i.e. with high prototypicality or highly prototypical features)—e.g. a dark-skinned African American person—are more negatively stereotyped than those with features that are closer to the characteristics of the majority group (low prototypicality or low prototypical features)—e.g. a light-skinned African American person (Maddox, 2004; Skinner and Nicolas, 2015; Wilkins *et al*., 2010). Although this research has mostly focused on race prototypicality in the US context, other external traits such as the Muslim headscarf can operate in practice as a physical trait, i.e. the Muslim headscarf has been perceived as an essential feature of women from specific ethnic minorities (Wiemers *et al*., 2024).

Consequently, based on the research on minority prototypicality bias, we expect ethnic minority individuals with highly prototypical features to have incentives to omit their photograph from their résumés to avoid being further penalized by employers. Likewise, we expect ethnic minority applicants with non-prototypical ethnic minority features to compensate for the negative stereotypes associated with their ethnic minority name by including a photograph that may (at least partially) weaken their association with a particular ethnic minority group.

We expect these incentives to be moderated by the country's social norms regarding the inclusion of headshots in résumés. In countries where this norm is strong, employers' penalty against minority applicants who do not include photographs is likely to be larger than the penalty against those who include headshots with high prototypical faces. Thus, applicants with high prototypical faces would not have incentives to omit their photograph from their résumé. In these contexts, we expect ethnic minority applicants who include photographs with low prototypical features to have the highest callback rate, followed by ethnic minority applicants whose résumés include a headshot with a highly prototypical appearance, and, at the bottom, ethnic minority applicants whose résumé does not include a photograph. In contrast, in countries where the social norm of including photographs on job applications is weak, we do not expect that failing to submit a photograph would be penalized by employers. Consequently, applicants with high prototypical appearance would have incentives to exclude photographs from their résumés, i.e. their ethnicity, race, or religious affiliation would not be immediately disclosed to employers, thus mitigating the negative affective responses that high prototypical traits may trigger. Conversely, applicants with low prototypicality will have incentives to include a picture to make their ethnic background—signalled by

their names—less salient for employers. Therefore, in contexts with a weak social norm regarding the inclusion of photographs, we expect minority applicants submitting a picture with low prototypical features to have the highest callback rate, followed by applicants who do not include pictures and, at the bottom, those including a photograph with highly prototypical features. The expected incentives for ethnic minority individuals to include a headshot based on their outward appearance are summarized in Table 1.

IV. Data and methods

We use data from the GEMM study, the first cross-national harmonized field experiment on discrimination in the labour market (Lancee *et al*., 2019, 2021).[1] The fieldwork took place between November 2016 and May 2018 in five European countries: Germany, the Netherlands, Norway, Spain, and the United Kingdom. The GEMM study employed an unpaired design whereby a single job application was submitted to each job vacancy in seven different occupational categories (cook, payroll clerk, receptionist, sales representative, software developer, shop assistant, and hairdresser). All fictitious candidates had four years of continuous work experience, and their ages ranged from 21 to 25 at the time of application, depending on the targeted occupation.

Here, we only use data for Germany and the Netherlands, two countries where the omission of headshot pictures in résumés was randomized by design. Based on previous correspondence audits and employers' surveys conducted in Germany (Kanning, 2016; Veit and Yemane, 2018; Weichselbaumer, 2017), it was assumed that the social norm associated with the inclusion of photographs in résumés in this country was strong, so only 10 per cent of résumés did not have headshot photographs. In contrast, in the Netherlands, it was assumed that the norm was weaker, as the inclusion of pictures in job applications had declined in popularity but had not disappeared (see, for instance, Aureus, 2024). Hence, photographs in the Netherlands were randomly included in half of the submitted applications.

The analytical sample includes majority group male and female job applicants—signalled with Dutch names in the Netherlands and German names in Germany—and ethnic minority applicants from the largest ethnic minorities in each country, namely the Turkish minority in Germany, and the Moroccan and Turkish minorities in the Netherlands.[2] Gender was randomly assigned, resulting in 51 per cent of job candidates in Germany and 50 per cent in the Netherlands having female names.

The GEMM study included additional experimental treatments that varied randomly across job applications. Along with the ethnicity and gender treatments, signalled mainly with names, there were treatments for candidates' religious affiliation, past job performance, and personality traits. In countries where attaching profile pictures to résumés was customary practice, a third of Moroccan and Turkish women were randomly assigned a headshot

Table 1: Ethnic minorities' incentives to include or omit a headshot picture from the résumé based on their outward appearance

Ethnic minority applicants' outward appearance	**Inclusion of headshots in résumés is a *strong* social norm (excluding photographs is penalized)**	**Inclusion of headshots in résumés is a *weak* social norm (excluding photographs is not penalized)**
Non-prototypical phenotype and appearance (close to majority group prototypical traits)	CBR of applications with photograph > CBR of applications without photograph	CBR of applications with photograph > CBR of applications without photograph
	Incentive to include it	**Incentive to include it**
Prototypical phenotype and appearance (different from majority group prototypical traits)	CBR of applications with photograph > CBR of applications without photograph	CBR of applications with photograph < CBR of applications without photograph
	Incentive to include it	**Incentive to exclude it**

Note: * CBR = callback rate.

[1] The GEMM study data is publicly available at https://doi.org/10.17026/dans-zrz-m9cm

[2] Ethnic background —signalled with forenames and surnames—was linked to applicants' country of birth for the majority group and ethnic minority applicants born abroad, or, among native born ethnic minority applicants, to the country of birth of their parents. To signal ethnic minority and majority membership, the most common first names and surnames in the country of origin at the time of applicants' birth year were chosen. In countries where data by birth year was not available for first names, websites listing countries' popular first names were used in the selection process.

wearing the Muslim headscarf. The photographs used for Dutch, German, Moroccan and Turkish applicants portrayed young men and women with different phenotypical traits. The range of phenotypes of candidates from the majority and minority groups were different[3] and are described in Table 2 (photographs included in the online appendix).

As stated in the last column of Table 2, we considered Turkish and Moroccan applicants with white phenotypes as having low (ethnic minority) prototypical features. Turkish and Moroccan applicants with brown or black skin were considered highly prototypical and distinct from the predominant majority group (white) phenotypes. Finally, we considered Turkish and Moroccan women wearing the Muslim headscarf as having high prototypicality, as they have been often stereotyped as traditional and *hijab* wearers (Wiemers *et al.*, 2024).

We used linear probability models (LPM) with robust standard errors. The outcome variable measures employers' responses, which were coded as positive ($y=1$) if the candidate received an invitation to an interview, a request to provide more information, or was shortlisted. Responses were coded as a negative callback ($y=0$) if the candidate received a rejection, just a confirmation of receipt, or no response at all. The number of applications submitted varied across occupations due to differences in labour demand between jobs, e.g. in the Netherlands, 24 per cent of applications were for vacancies for cooks, compared to only 5 per cent for vacancies for hairdressers. Thus, all models use occupation weights, which we created by dividing the average number of applications per occupation (numerator) by the number of applications submitted to each occupation (denominator) in each country $\left(\frac{\text{total no. applications}/\text{no. occupations (7)}}{\text{no. applications per occupation}}\right)$ That way, we ensure that each of the seven occupations has the same weight in the estimation, thus preventing the results from being driven by the occupations with more observations.

V. Assessing the strength of the social norm of including photographs in résumés in Germany and the Netherlands

In the research design of the GEMM study, it was assumed that the norm of including photographs in job applications was stronger in Germany than in the Netherlands. However, to validate this assumption, it is necessary to investigate employers' responses towards majority group applicants who submitted applications with and without photographs. If the norm is indeed strong in Germany, we anticipate a disadvantage for the majority group applicants who omitted headshot pictures compared to those who included them. Conversely, if the norm is weak in the Netherlands, we would expect no systematic differences in employers' responses towards applicants based on whether they included a photograph or not. We further examine whether the social norm regarding the inclusion of photographs in résumés varies depending on the gender of the applicants, particularly in Germany, where we anticipate that applications lacking photographs may face penalties. Some previous studies have shown that the

Table 2: Phenotypes used for majority and minority group applicants

Phenotypes	Used by majority group applicants	Used by Moroccan applicants	Used by Turkish applicants	Degree of prototypicality for Moroccan and Turkish applicants
Phenotype 1: White, with blonde hair and blue eyes	Yes	No	No	Low
Phenotype 2: White, with light brown hair and green eyes.	Yes	No	Yes	Low
Phenotype 3: White, with darker skin, dark brown hair and eyes.	Yes	Yes	Yes	Low
Phenotype 4: Brown, dark brown hair and eyes (Arab features)	Yes	Yes	Yes	High
Phenotype 5: Brown, dark brown hair and eyes (Southeast Asian features)	Yes	No	Yes	High
Phenotype 6: Black, black hair, dark brown eyes	No	Yes	No	High

[3] The GEMM study experimental design maximized the within-ethnicity phenotypic variation, but the two treatments (ethnicity and phenotype) were not orthogonal. An online survey with a convenient student sample was conducted to determine which of the different phenotypes could be plausible among the population from different countries of origin.

beauty premium in the labour market is gendered (Kukkonen *et al*., 2024), suggesting that the relevance of submitting a photograph could differ for male and female applicants. It is therefore plausible that the inclusion of a photograph might be more advantageous for male candidates than for female candidates, or vice versa.

In order to answer this question, we estimated the following LPM in each country:

$$Y_i = \beta_0 + \beta_2 Photo_i + \beta_3 Gender_i + \beta_4 (Photo\ x\ Gender)_i + e_i \quad (1)$$

Y_i is the dependent variable for observation *i*. β_0 is the intercept; *Photo* is the categorical variable indicating whether a photograph was included in the application or not; *Gender* indicates whether job applicants had a male or female name. Finally, *Photo x Gender* represents the interaction term. The estimated linear predictions based on the coefficients of Equation 1 are presented in Figure 1. The model coefficients can be found in the online appendix.

Contrary to our expectations, Dutch women who included a headshot picture in their application had a significantly higher callback rate than those who did not. The difference between the predicted callback rates of the two groups is statistically significant (p-value = 0.003). Similarly, Dutch men who included a photograph had a higher callback rate than those who omitted it, although the difference between the two groups did not reach statistical significance (p-value>0.05). It is possible that the premium given to Dutch men who attached photographs was smaller compared to women who did so, but we would have required a larger sample size to detect that effect. In contrast, in Germany, there were no statistically significant differences in callback rates between applicants who submitted a photograph and those who did not, regardless of their gender. Note, however, that the callback rate of female applicants who did not include a photograph was higher than those who included it, although the difference between the two groups is not statistically significant (p-value = 0.093). The confidence intervals of the estimated callback rates for applicants who did not submit photographs in Germany are wider, likely due to the smaller sample size (only 12 per cent of applications were submitted without photographs).

To summarize, the inclusion of a photograph significantly enhances the employability prospects for female applicants in the Netherlands. However, for Dutch male applicants and applicants of both genders in Germany, there does not appear to be a premium or penalty associated with including a photograph in job applications. Thus, we conclude that the norm of including photographs in résumés may not have been as strong in Germany as initially imagined, whereas it may have been stronger than expected in the Dutch context, at least for female candidates.

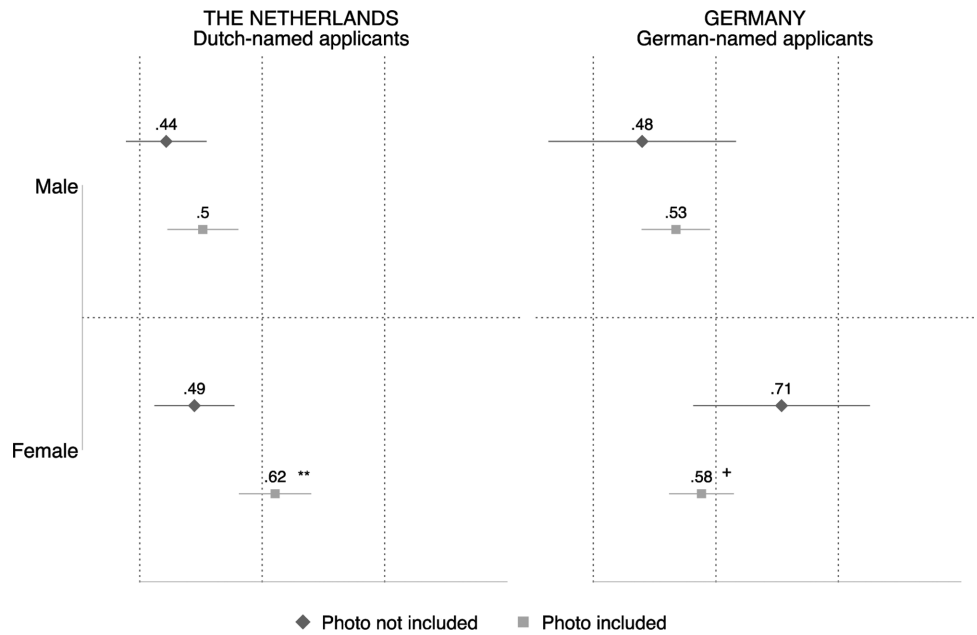

Figure 1: Probability of receiving a positive callback from employers in Germany and the Netherlands. Male and female majority group applicants. Linear predictions.

Notes: Germany n=781; the Netherlands n= 1,031. Different callback rate relative to baseline (no photograph): ** p<0.005 * p<0.05 +p<0.10.

VI. Minorities' incentives to include or omit headshot pictures from résumés: differences based on individuals' phenotypic traits in the Netherlands and Germany

To answer this question, we estimated the following LPM separately for male and female applicants in each country:

$$Y_i = \beta_0 + \beta_1 Name_i + \beta_2 Phenotype_i + \beta_3 Name_i + \beta_4 (Name\ x\ Phenotype)_i + e_i \quad (2)$$

Y_i is the dependent variable for observation *i*. β_0 is the intercept; *Phenotype* is the categorical variable indicating whether a photograph was included in the application and, if so, the phenotype depicted in the picture; *Name* indicates whether job applicants had a majority group (Dutch/German) or minority (Turkish/Moroccan) name. Finally, *Name x Phenotype* represents the interaction terms. We focus on specific linear predictions based on the coefficients of equation 2, which are included in the online appendix.

In the Netherlands (Figure 2), failing to attach a headshot picture was penalized for Turkish female applicants, but only when comparing them to female applicants submitting a picture with a white phenotype (phenotypes 1, 2 and 3). That is, the differences in the estimated callback rates between Turkish women who omitted their headshot picture and those who submitted a photograph with low phenotypicality (white phenotypes 2 and 3) are statistically significant (p-value< 0.000). The callback rate of Turkish women who applied for a job with a high phenotypical headshot picture (brown phenotypes 4 and 5) was also slightly higher than that of Turkish women who omitted their picture, although the differences between the two groups are not statistically significant (p-value = 0.078). As for Moroccan women, we did not find evidence that including a headshot picture increased or decreased their callback rate with respect to those who did not include it. For both majority and (Turkish and Moroccan) minority group men, we did not detect statistically significant differences between those who omitted their headshot picture from the résumé and those who included photographs conveying different phenotypes.

In Germany (Figure 3), we did not detect statistically significant differences between ethnic minority candidates who included photographs in their résumés and those who did not. Turkish women submitting application pictures with low phenotypical traits (white phenotypes 2 and 3) seemed to be advantaged relative to those who did not include a headshot photograph in their résumé, but the difference between the two groups' callback rates is not statistically significant at conventional levels, likely due to the small sample size. No statistically significant differences are detected between applications of Turkish men with and without photographs. It is also worth noting that

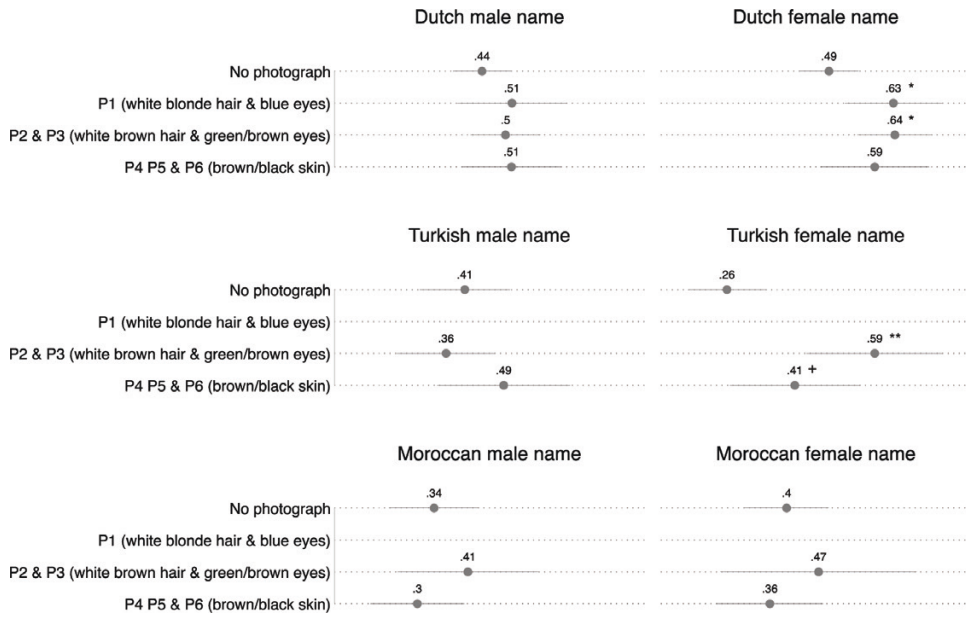

Figure 2: Probability of receiving a positive callback from employers in the Netherlands, by applicants' name and phenotype. Linear predictions

Notes: n= 1,818. Applicants wearing the Muslim headscarf are excluded.

Different callback rate relative to baseline (no photograph): ** p<0.005 * p<0.05 +p<0.10.

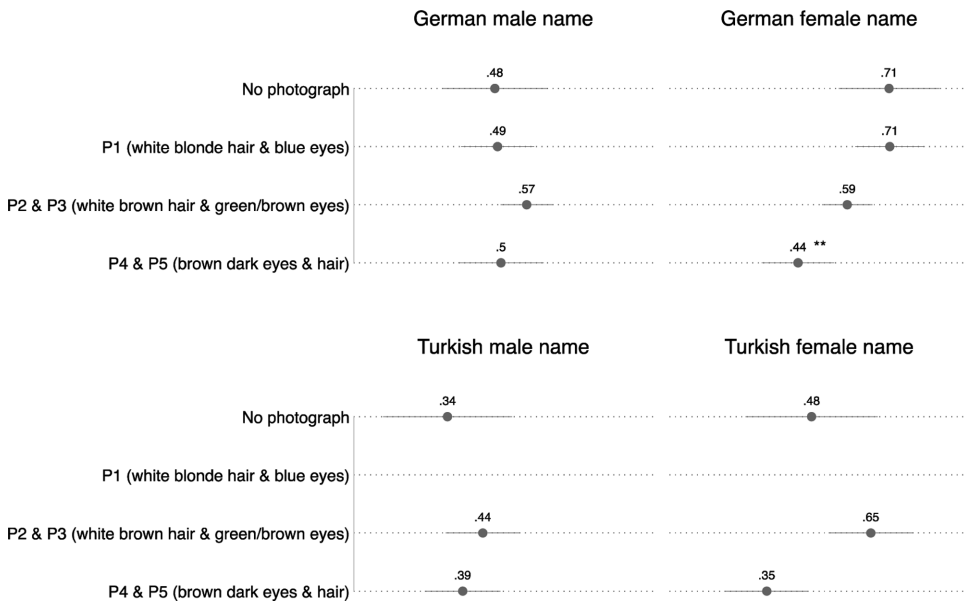

Figure 3: Probability of receiving a positive callback from employers in Germany, by applicants' name and phenotype. Linear predictions.
Notes: n= 1,129. Applicants wearing the Muslim headscarf are excluded. Different callback rate relative to baseline (no photograph): ** p<0.005 * p<0.05 +p<0.10.

Turkish women with white phenotypes had a statistically significant higher callback rate than those whose picture was highly prototypical (brown phenotypes 4 and 5) (p-value < 0.000). Likewise, brown skin German women (phenotypes 4 and 5) had a statistically significant lower callback rate than those with prototypical German traits (white phenotypes 1, 2 and 3) (p-value = 0.004). This suggests that dark skin tones were penalized among female applicants, even among those with German names (see Polavieja *et al*. 2023).

VII. Minority women's incentives to include or omit headshot pictures from résumés when they wear the Muslim headscarf

We estimated the following LPM for the sample of ethnic minority women with Turkish and Moroccan names in the Netherlands and Germany:

$$Y_i = \beta_0 + \beta_1 Name_i + \beta_2 Photograph_i + e_i \quad (The\ Netherlands) \tag{3}$$

$$Y_i = \beta_0 + \beta_2 Photograph_i + e_i \quad (\text{Germany}) \tag{4}$$

Y_i is the dependent variable (callback rate) for observation *i*. β_0 is the intercept; *Name* indicates whether job applicants had a Turkish or a Moroccan name (for the Dutch sample); and *Photograph* is the treatment variable which takes the following values: (1) no photograph; (2) headshot photograph with low phenotypicality and no Muslim headscarf; (3) photograph with high phenotypicality and no Muslim headscarf, (4) photograph with low phenotypicality wearing the Muslim headscarf; and (5) photograph with high phenotypicality wearing the Muslim headscarf. Figure 4 depicts the linear predictions based on the coefficients of the model above, which are included in the online appendix.

In the Netherlands, minority women with low prototypical faces (white phenotypes 2 and 3) who did not wear the Muslim headscarf had a higher callback rate than minority women who did not include a headshot picture. They also had a higher callback rate than those wearing a Muslim headscarf in their headshot picture, regardless of phenotype (p-value = 0.007). These differences in callback rates are not found for minority women wearing the Muslim headscarf with high prototypical faces (phenotypes 4, 5 and 6), whose callback rate is not different from minority women who did not include a photograph in their job application. Hence, Turkish and Moroccan women in the Netherlands would only have an incentive to include a picture in their application if they do not wear the Muslim headscarf and have a white phenotype. As a result, veiled minority women would not have an incentive

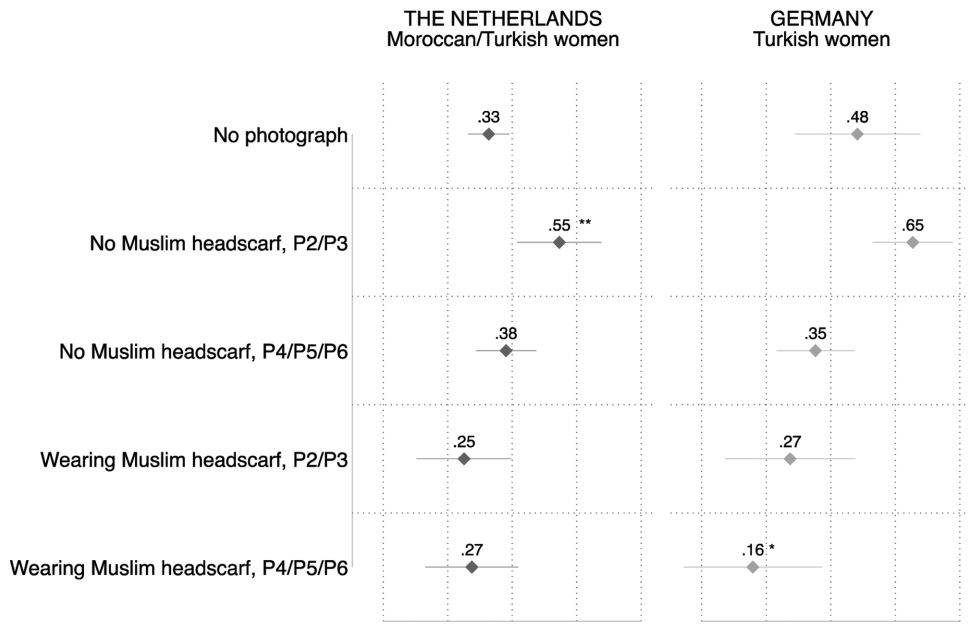

Figure 4: Probability of receiving a positive callback in Germany and the Netherlands (ethnic minority women). Linear predictions.
The Netherlands n=438; Germany n=187
Different callback rate relative to baseline (no photograph): ** p<0.005 * p<0.05 +p<0.10

to omit their picture, as they would receive similar treatment from employers regardless of whether they include it or not.

In Germany, non-White minority women (phenotypes 4 and 5) who submitted a picture wearing the Muslim headscarf would have incentives to omit their picture from their application, as they were penalized compared to minority women who did not include a headshot picture in their résumé—the difference between the two groups is statistically significant, with a p-value of 0.028.

VIII. Discussion of results

The GEMM study did not measure ethnic minorities' expectations regarding employers' behaviour during the screening process. However, the data enabled us to examine the incentives that minority men and women may have to include or exclude headshot pictures from job applications based on observed employers' responses towards candidates with different outward appearances.

The results show that, contrary to our initial assumptions, omitting headshot pictures from résumés was on average not penalized in Germany, whereas in the Netherlands, majority group women who attached pictures received a premium compared to those who did not. Assuming that we can rely on employers' behaviour towards majority group applicants to identify their true preferences regarding this practice, we concluded that omitting photographs from job applications was, unexpectedly, only consequential for women in the Netherlands. In this regard, our analysis highlights the challenges of predicting employers' biases on this matter, as their true preferences may not be evident for job applicants. A limitation of our approach is that we did not infer the strength of the social norm based on the prevalence of including photographs in job applications in the two countries. It is possible that this practice was not widespread in the Netherlands at the time of the fieldwork, i.e. candidates submitting photographs could well have been a minority in the Dutch labour market. Regardless of how common this practice was, employers still valued female applicants who included a photograph in their résumés. As for Dutch men, either the premium was smaller than for women and we were not able to detect it in the empirical analysis, or submitting photographs was inconsequential for them (i.e. weak social norm).

We expected that, in both Germany and the Netherlands, minority applicants submitting photographs with low prototypical features (white phenotypes) would have a higher callback rate than those who did not submit a photograph and, hence, an incentive to always include photographs. This was only confirmed for Turkish women

in the Netherlands, but not for Moroccan women, for whom we did not detect a statistically significant advantage of including a low prototypical picture relative to submitting their applications without a photograph. In Germany, we were not able to detect statistically significant differences between Turkish applicants who omitted their photograph and those who included a picture with low prototypical features, despite the latter group having a higher callback rate. This lack of statistical significance is possibly due to the small sample size. Likewise, in both Germany and the Netherlands, we did not detect a statistically significant difference in callback rates between minority men including a photograph with low prototypicality in their résumé and those who did not include any photograph.

With regard to minority applicants with high prototypicality (with brown and black phenotypes, or who wore the Muslim headscarf), we hypothesized that they would have the incentive to submit a photograph in contexts where the social norm of including pictures in job applications was strong. In contrast, we expected that they would have an incentive to omit photographs where the social norm was weak. Consequently, given the positive effect of attaching photographs for majority group women in the Dutch context, we expected ethnic minority women with high prototypical features would have incentives to include their pictures. However, this was not confirmed. We could not detect statistically significant differences at $p<0.05$ between Turkish women submitting a photograph with brown phenotypes and those who omitted a headshot photograph (p-value = 0.078). Similarly, for Moroccan women and women wearing the Muslim veil, we did not detect a statistically significant advantage (or disadvantage) relative to those omitting photographs. We also expected that highly prototypical minority men in Germany and the Netherlands and highly prototypical minority women in Germany would be better off omitting their pictures from their résumé rather than including them. This expectation was only confirmed for women wearing the Muslim headscarf in Germany, who had a lower callback rate than minority women who did not include any picture in the job application.

To summarize, Turkish women in the Netherlands who did not wear the Muslim veil were likely better off submitting their picture rather than excluding it. The advantage associated to including a picture was higher for Turkish women with low prototypicality (white phenotypes) than for those with high prototypicality brown phenotypes, for which the advantage was only marginally significant. The premium associated with submitting a picture for female applicants in the Netherlands disappeared when minority women wore the Muslim headscarf. In that case, including a photograph in the résumé would not confer them any advantage relative to omitting it. For unveiled Moroccan women and Moroccan/Turkish men in the Netherlands, there was also no discernible advantage or disadvantage associated with the inclusion of photographs. In Germany, we could not detect any significant advantage or disadvantage of including pictures with low or high prototypicality, although we may have been able to detect those differences for Turkish women if we had a larger sample size. For instance, majority group women with brown phenotypes had a lower callback rate relative to those who did not include any picture in Germany, which suggests that phenotypes mattered for employers when assessing female candidates. Finally, in the German context women wearing the Turkish veil would be better off excluding their picture from their résumé.

Overall, the results show that, depending on the context, minority women had incentives to include or exclude their photographs from job applications, but not men. In other words, outward appearance seemed to be more consequential for female than for male candidates. Notably, minority women with high prototypical appearance (with brown phenotypes or wearing the Muslim headscarf) would only have incentives to engage in résumé whitening—i.e. omitting photographs from job applications—in the German context, particularly if they wore the Muslim headscarf. Note, however, that a similar correspondence audit conducted in the German context between 2014 and 2016 (see Veit and Yemane, 2018) found that some job application systems specifically asked prospective applicants to upload their photographs, which suggests that submitting at application without a photograph may not have been an option, at least ten years ago.

The findings underscore the heterogeneity of factors that ethnic minority women may weigh when deciding whether to include a photograph in their résumé. Considerations such as prevailing social norms regarding photograph inclusion, their ethnicity, or the degree of phenotypicality of their appearance can all significantly influence employer behaviour. Minority job applicants may struggle to gauge the level of discrimination they might face in various contexts, and to decide how best to apply for specific jobs (e.g. including or excluding photographs). A limitation of our study is that it did not allow us to evaluate the effect of other types of résumé whitening, e.g. using a majority group name while maintaining an ethnic minority surname (see Fossati *et al.*, 2020). Another limitation was the small sample size of applications submitted without photographs in the German correspondence audit, a factor that affected our ability to detect statistically significant differences between groups.

IX. What to do next? Interventions to eliminate discrimination in recruitment practices

The results discussed above indicate that, in general, ethnic minorities may find it challenging to accurately assess the strength of the social norm regarding the inclusion of pictures in résumés and the potential penalties for failing to do so. This suggests that omitting headshot pictures from résumés as a strategy to reduce discrimination based on outward appearance may be counterproductive if minorities mistakenly believe that employers do not penalize applicants who do not include photographs.

A relatively easy intervention to minimize the effect of employers' reactions towards different ethnic, racial or religious cues, and to alleviate minorities from the stress of deciding whether to include or omit a photograph, would involve the prohibition of photographs in job applications or the adoption of blind recruitment or full anonymization policies. Importantly, blind recruitment not only entails the exclusion of photographs from résumés, a practice already enforced in countries such as the UK or the US, but also encompasses the omission of identifying information from job applications, such as names, places of residence, and even educational institutions. Indeed, names are relied upon by employers to infer applicants' gender and ethnicity in contexts where photographs are not allowed, e.g. recent correspondence audits conducted in the UK showed employers' discrimination against Pakistani and Black Caribbean (Heath and Di Stasio, 2019; Rovira, 2023). While not yet a widespread practice, some organizations, particularly in the public sector, have implemented blind recruitment in the first stages of the hiring process, e.g. the UK Civil Service (Manzoni, 2015) and the British Library (Vercida, 2019).

Several field experiments have been conducted to measure the effect of anonymization policies on ethnic discrimination rates in the recruitment process. The results show that the differences in callback rates between comparable minority and majority group candidates disappear when anonymous job applications are introduced (Bøg and Kranendonk, 2011; Åslund and Skans, 2012; Blommaert and Coenders, 2023), but see Behaghel *et al*. (2015). This is not surprising: if application documents preserve anonymity effectively, discrimination becomes impossible. Krause *et al*. (2012) and Blommaert and Coenders (2023) showed how the absence of discrimination depends on the anonymization process, and they suggest that standardized forms seem to be the best method.

However, full anonymized recruitment should be carefully executed (Rinne, 2018; McGinnity *et al*., 2021). First, limiting the information that recruiters can see on applicants' minority backgrounds may hinder other positive measures to increase the ethnic diversity of the workforce and can limit the scope for affirmative action. For instance, Behaghel *et al*. (2015) found that ethnic minorities received fewer invitations to continue the recruitment process when blind recruitment was implemented. The authors explained the result based on the self-selection of firms that voluntarily took part in the study, i.e. these companies had a higher propensity to hire more minority candidates than other non-participating firms, likely because they were already committed to increasing the diversity of their workforce. The complete anonymization of résumés did not allow them to use ethnicity as a tie-breaking method. Second, recruiters may interpret context-specific information disadvantageously if the candidate's identity or personal history is unknown (Rinne, 2018; McGinnity *et al*., 2021), e.g. recruiters may be less inclined to penalize candidates for their unemployment spells if they are from an ethnic minority, as those candidates often face more adverse labour market conditions. Behaghel *et al*. (2015) also found that negative signals in application documents, such as interrupted labour market histories, were judged more negatively by employers when applicants' minority group status was not visible to them.

Blind recruitment policies serve as a valuable tool in mitigating ethnic discrimination during the hiring process, yet they have limitations and should not be considered a universal solution for ethnic discrimination and inequalities. First, blind recruitment policies may simply postpone discrimination to later stages of the hiring process. Indeed, the research by Bøg and Kranendonk (2011) in the Netherlands and Åslund and Skans (2012) in Sweden showed that anonymous hiring increased the likelihood of ethnic minorities securing an interview but did not translate into an increased probability of receiving a final job offer at the second stage. This is a particular concern for experimental audit studies, which only focus on the first stage of the recruitment process (Quillian *et al*., 2020). Second, anonymization only works for openly advertised job openings. Many job opportunities are not widely advertised, and people learn about them through personal networks and recommendations (Granovetter, 1995; Pellizzari, 2010). In these situations, anonymization policies would have little effect. Third, blind recruitment cannot correct on its own the results of cumulative discrimination over the life course that puts some ethnic minorities in a disadvantaged position when applying for jobs (Blank, 2005). Discrimination is prevalent for minority applicants in the education system (D'hondt *et al*., 2021), housing, and financial markets (Pager and Shepherd, 2008), thus impacting school choice and educational trajectories (McGinnity *et al*., 2021). Consequently, the CVs of minority applicants could appear, on average, less competitive than those from the majority group. In such cases, anonymization will not correct the effects of past discrimination and inequalities. Similarly, if characteristics that are more likely to be possessed by majority groups are more valued, e.g., assertiveness by White males (Fukuyama and Greenfield, 1983), then inequality dynamics will remain unchanged regardless of anonymization.

Full anonymization policies show promise in mitigating direct discrimination during the initial stages of recruitment. Additionally, they can make ethnic minorities more comfortable applying for jobs, as they would not have to engage in strategic résumé whitening. However, the efficacy of these policies depends on their careful execution, with potential drawbacks including the inability to implement diversity hiring and the risk of postponing discrimination to later stages in the hiring process. Moreover, these policies can only mitigate direct discrimination at a single point in time and are unable to tackle biases ingrained in job evaluation criteria. Overall, these limitations underscore the necessity for a holistic approach to anti-discrimination policies that go beyond anonymized job recruitment.

The code and data to reproduce our analysis is available at https://osf.io/h68dr/

Supplementary data

Supplementary data is available at *Oxford Review of Economic Policy Journal* online.

References

Åslund, O., and Skans, O. N. (2012), 'Do Anonymous Job Application Procedures Level the Playing Field?', *ILR Review*, **65**(1), 82–107.

Aureus (2024), 'Do I Need a Photo on my Resume?', Study Association Aureus, retrieved from https://www.aureus.nl/need-photo-resume, 11 March 2024.

Behaghel, L., Crépon, B., and Le Barbanchon, T. (2015), 'Unintended Effects of Anonymous Résumés'm *American Economic Journal: Applied Economics*, **7**(3), 1–27.

Bertrand, M., and Duflo, E. (2017), 'Field Experiments on Discrimination', in A. Banerjee and E. Duflo (eds), *Handbook of Economic Field Experiments* (Vol. **1**), North Holland, 309–93.

— Mullainathan, S. (2004), 'Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination', *American Economic Review*, **94**(4), 991–1013.

Blank, R. M. (2005), 'Tracing the Economic Impact of Cumulative Discrimination', *American Economic Review*, **95**(2), 99–103.

Blommaert, L., and Coenders, M. (2023), 'The Effects of and Support for Anonymous Job Application Procedures: Evidence from a Large-scale, Multi-faceted Study in the Netherlands', *Journal of Ethnic and Migration Studies*, DOI: 10.1080/1369183X.2023.2282385.

Bøg, M., and Kranendonk, E. (2011), 'Labor Market Discrimination of Minorities? Yes, but not in Job Offers', MPRA Paper N° 33332.

Daniel, W. W. (1968), *Racial Discrimination in England*, London, Penguin.

Di Stasio, V., Lancee, B., Veit, S., and Yemane, R. (2021), 'Muslim by Default or Religious Discrimination? Results from a Cross-national Field Experiment on Hiring Discrimination', *Journal of Ethnic and Migration Studies*, **47**(6), 1305–26.

Dixon, A. R., and Telles, E. E. (2017), 'Skin Color and Colorism: Global Research, Concepts, and Measurement', *Annual Review of Sociology*, **43**(1), 405–24.

D'hondt, F., Maene, C., Vervaet, R., Van Houtte, M., and Stevens, P. A. J. (2021), 'Ethnic Discrimination in Secondary Education: Does the Solution Lie in Multicultural Education and the Ethnic School Composition?', *Social Psychology of Education*, **24**(5), 1231–58.

Equality and Human Rights Commission (2011), *Equality Act 2010 Code of Practice. Employment. Statutory Code of Practice*, London, Equality and Human Rights Commission.

Fernández-Reino, M., Di Stasio, V., and Veit, S. (2023), 'Discrimination Unveiled: A Field Experiment on the Barriers Faced by Muslim Women in Germany, the Netherlands, and Spain', *European Sociological Review*, **39**(3), 479–97.

Fossati, F., Liechti, F., and Auer, D. (2020), 'Can Signaling Assimilation Mitigate Hiring Discrimination? Evidence from a Survey Experiment', *Research in Social Stratification and Mobility*, **65**, 100462.

Fukuyama, M. A., and Greenfield, T. K. (1983), 'Dimensions of Assertiveness in an Asian-American Student Population', *Journal of Counseling Psychology*, **30**(3), 429–32.

Gaddis, S. M. (2019), *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, Berlin, Springer.

Granovetter, M. (1995), *Getting A Job: A Study of Contacts and Careers*, Chicago, IL, University of Chicago Press.

Heath, A., and Di Stasio, V. (2019), 'Racial Discrimination in Britain, 1969–2017: A Meta-analysis of Field Experiments on Racial Discrimination in the British Labour Market', *The British Journal of Sociology*, **70**(5), 1774–98.

Kang, S. K., DeCelles, K. A., Tilcsik, A., and Jun, S. (2016), 'Whitened Résumés', *Administrative Science Quarterly*, **61**(3), 469–502.

Kanning, U. P. (2016), 'Über die Sichtung von Bewerbungsunterlagen in der Praxis der Personalauswahl', *Zeitschrift Für Arbeits- Und Organisationspsychologie A&O*, **60**(1), 18–32.

Krause, A., Rinne, U., and Zimmermann, K. F. (2012), 'Anonymous Job Applications in Europe', *IZA Journal of European Labor Studies*, **1**(1), 5.

Kukkonen, I., Pajunen, T., Sarpila, O., and Åberg, E. (2024), 'Is Beauty-based Inequality Gendered? A Systematic Review of Gender Differences in Socioeconomic Outcomes of Physical Attractiveness in Labor Markets', *European Societies*, **26**(1), 117–48.

Lancee, B., Birkelund, G., Coenders, M., Di Stasio, V., Fernández Reino, M., Heath, A., Koopmans, R., Larsen, E. N., Polavieja, J., Ramos, M., Thijssen, L., Veit, V., and Yemane, R. (2021), *The GEMM Study: A Cross-national Harmonized Field Experiment on Hiring Discrimination*.

Lancee, B., Zwier, D., Birkelund, G., Coenders, M., Di Stasio, V., Fernández Reino, M., Heath, A., Koopmans, R., Larsen, E., Polavieja, J., Ramos, M., Thijssen, L., Veit, S., and Yemane, R. (2019), *The GEMM Study: A Cross-national Harmonized Field Experiment on Labour Market Discrimination: Technical Report*.

Lippens, L., Vermeiren, S., and Baert, S. (2023), 'The State of Hiring Discrimination: A Meta-analysis of (Almost) All Recent Correspondence Experiments', *European Economic Review*, **151**, 104315.

Livingston, R. W., and Brewer, M. B. (2002), 'What are We Really Priming? Cue-based versus Category-based Processing of Facial Stimuli', *Journal of Personality and Social Psychology*, **82**(1), 5–18.

McGinnity, F., Quinn, E., McCullough, E., Enright, S., and Curristan, S. (2021), *Measures to Combat Racial Discrimination and Promote Diversity in the Labour Market: A Review of Evidence*, Economic and Social Research Institute.

Maddox, K. B. (2004), 'Perspectives on Racial Phenotypicality Bias', *Personality and Social Psychology Review*, **8**(4), 383–401.

Manzoni, J. (2015), 'Name-blind Recruitment—A Commitment to Diversity', GOV.UK Civil Service Blog, 5 November, https://civilservice.blog.gov.uk/2015/11/05/name-blind-recruitment-a-commitment-to-diversity/.

Marira, T. D., and Mitra, P. (2013), 'Colorism: Ubiquitous Yet Understudied', *Industrial and Organizational Psychology*, **6**(1), 103–7.

Pager, D. (2007), 'The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future', *The ANNALS of the American Academy of Political and Social Science*, **609**(1), 104–33.

— Karafin, D. (2009), 'Bayesian Bigot? Statistical Discrimination, Stereotypes, and Employer Decision Making', *Annual Review of Sociology*, **621**(1), 70–93.

— Shepherd, H. (2008), 'The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets', *Annual Review of Sociology*, **34**(1), 181–209.

— Western, B., and Bonikowski, B. (2009), 'Discrimination in a Low-wage Labor Market: A Field Experiment', *American Sociological Review*, **74**(5), 777–99.

Pellizzari, M. (2010), 'Do Friends and Relatives Really Help in Getting a Good Job?', *ILR Review*, **63**(3), 494–510.

Polavieja, J. G., Lancee, B., Ramos, M., Veit, S., and Yemane, R. (2023), 'In your Face: A Comparative Field Experiment on Racial Discrimination in Europe', *Socio-Economic Review*, **21**(3), 1551–78.

Quillian, L., and Lee, J. J. (2023), 'Trends in Racial and Ethnic Discrimination in Hiring in Six Western Countries', *Proceedings of the National Academy of Sciences*, **120**(6).

— Midtbøen, A. H. (2021), 'Comparative Perspectives on Racial Discrimination in Hiring: The Rise of Field Experiments', *Annual Review of Sociology*, **47**(1), 391–415.

— Lee, J. J., and Oliver, M. (2020), 'Evidence from Field Experiments in Hiring Shows Substantial Additional Racial Discrimination after the Callback', *Social Forces*, **99**(2), 732–59.

— Heath, A., Pager, D., Midtbøen, A., Fleischmann, F., and Hexel, O. (2019), 'Do Some Countries Discriminate More than Others? Evidence from 97 Field Experiments of Racial Discrimination in Hiring', *Sociological Science*, **6**, 467–96.

Rich, J. (2014), 'What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted Since 2000', IZA Discussion Paper, No. 8584.

— (2018), 'Do Photos Help or Hinder Field Experiments of Discrimination?', *International Journal of Manpower*, **39**(4), 502–18.

Rinne, U. (2018), 'Anonymous Job Applications and Hiring Discrimination', *IZA World of Labor*.

Rooth, D.-O. (2009), 'Obesity, Attractiveness, and Differential Treatment in Hiring', *Journal of Human Resources*, **44**(3), 710–35.

Rovira, M. (2023), 'Invisible Stripes? A Field Experiment on the Disclosure of a Criminal Record in the British Labour Market and the Potential Effects of Introducing Ban-the-Box Policies', *The British Journal of Criminology*, DOI: 10.1093/bjc/azad063.

Ruedin, D., and Van Belle, E. (2023), 'The Extent of Résumé Whitening', *Sociological Research Online*, **28**(3), 858–69.

Ruffle, B. J., and Shtudiner, Z. (2015), 'Are Good-looking People More Employable?', *Management Science*, **61**(8), 1760–76.

Simon, P. (2017), 'The Failure of the Importation of Ethno-racial Statistics in Europe: Debates and Controversies', *Ethnic and Racial Studies*, **40**(13), 2326–32.

Skinner, A. L., and Nicolas, G. (2015), 'Looking Black or Looking Back? Using Phenotype and Ancestry to Make Racial Categorizations', *Journal of Experimental Social Psychology*, **57**, 55–63.

Thijssen, L., van Tubergen, F., Coenders, M., Hellpap, R., and Jak, S. (2022), 'Discrimination of Black and Muslim Minority Groups in Western Societies: Evidence from a Meta-analysis of Field Experiments', *International Migration Review*, **56**(3), 843–80.

Veit, S., and Yemane, R. (2018), 'The ADIS Study: A Large-scale Correspondence Test on Labor Market Discrimination in Germany—Technical Report', https://hdl.handle.net/10419/178706

Vercida (2019), 'Anonymous Recruitment Introduced to Reduce Unconscious Bias', https://www.vercida.com/uk/articles/anonymous-recruitment-british-library.

Weichselbaumer, D. (2017), 'Discrimination Against Migrant Job Applicants in Austria: An Experimental Study', *German Economic Review*, **18**(2), 237–65.

Wiemers, S. A., Stasio, V. Di, and Veit, S. (2024), 'Stereotypes about Muslims in the Netherlands: An Intersectional Approach', *Social Psychology Quarterly*, https://doi.org/10.1177/01902725231219688.

Wilkins, C. L., Kaiser, C. R., and Rieck, H. (2010), 'Detecting Racial Identification: The Role of Phenotypic Prototypicality', *Journal of Experimental Social Psychology*, **46**(6), 1029–34.

Zschirnt, E., and Ruedin, D. (2016), 'Ethnic Discrimination in Hiring Decisions: A Meta-analysis of Correspondence Tests 1990–2015', *Journal of Ethnic and Migration Studies*, **42**(7), 1115–34.