

Can skills alleviate ethnic discrimination in hiring? Evidence from the German apprenticeship market.

Alina Shirshikova^{*a}, Frank Cörvers^a, Raymond Montizaan^{a,c}, Harald Pfeifer^{a,b,c}, and Caroline Wehner^{b,c}

^a*Research Centre for Education and the Labour Market (ROA), Maastricht University,
Netherlands*

^b*Federal Institute for Vocational Education and Training (BIBB), Germany*

^c*Institute of Labor Economics (IZA), Germany*

November, 2023

Abstract

This study investigates whether the social, digital, and analytical skills of job applicants, as well as their knowledge about the profession, can mitigate ethnic disparities in entry-level positions in the labour market. We conducted a survey experiment among a large, nationally representative sample of German firms that hire apprentices. We asked recruiters to evaluate the probability of inviting fictitious applicants to a job interview based on randomised characteristics, including ethnicity, skill quality, gender, time of residence, and education level. Our results show heterogeneous effects of skills on ethnic discrimination. While social skills help alleviate discrimination, our results indicate that discrimination intensifies at higher levels of knowledge about the profession, implying greater disparities due to ethnic discrimination at the top of the skill distribution. We also found that the effect of skills differs depending on the ethnicity of the applicant.

JEL classification: I24, J15, J23, J24, J71

Keywords: ethnic discrimination, hiring, skills, vocational education and training, apprenticeships

*Corresponding author, email: a.shirshikova@maastrichtuniversity.nl

We are grateful to Heike Solga, Dominik Becker and Saena Chakkar for their insightful comments, to the Federal Institute for Vocational Education and Training (BIBB) which made the study possible, and to the audiences at the Workshop on Survey Experiments in Migration and Integration Research 2021, EALE 2023, the IAB/LASER-Workshop on Survey Experiments in Labor Market and Regional Research 2023, the MORSE Conference 2023 and the LIMES Conference 2023 for their useful questions, discussions, and encouragement.

This work was supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 847596.

1 Introduction

Labour shortages constitute a notable problem in many European countries (European Labour Authority, 2022). To address this issue, the European Commission aims to facilitate labour mobility between EU Member States and to attract people from non-EU countries, including by strengthening learning opportunities (European Commission, 2022, page 4). Meanwhile, a large strand of the literature shows that employment discrimination on the basis of ethnicity and national origin remains a persistent social problem (Lippens et al., 2023; Neumark, 2018; Quillian et al., 2017) that may hinder mobility and lead to sub-optimal allocation of labour resources. Investigating factors that have the potential to moderate discrimination in hiring is valuable from both research and policy perspectives.

In recent years, the emerging literature on decision-making in hiring has documented the crucial role of skills in the hiring process. Several studies show that the job applicant's cognitive, non-cognitive and occupation-specific skills greatly affect the probability of being selected for a job, especially for entry-level positions (Deming, 2017; Humburg & van der Velden, 2015; Piopiunik et al., 2020). While existing research on ethnic discrimination mostly focuses on studying the magnitude and nature of discriminatory biases (Baert, 2018; Bertrand & Duflo, 2017; Quillian et al., 2020), as well as its dependence on firm and vacancy characteristics (Kline et al., 2022; Kline & Walters, 2021), an aspect that has received less attention by the literature concerns factors that can mitigate discriminatory behaviour. In particular, little is known about whether and to what extent skills can alleviate disparate treatment of applicants with ethnic backgrounds.

In this study, we empirically investigate how social skills, knowledge about the profession, and the digital and analytical skills of the applicants affect labour-market entry and whether these skills can mitigate disparities in the probability of an invitation to a job interview across different ethnicities.¹ To do so, we conducted a factorial survey

¹We define *social skills* as an ability to work in teams, communicate and cooperate effectively. *Knowledge about the profession* reflects the awareness of the job requirements and job-related experience. *Digital skills* stand for the use of relevant software and programming skills. *Analytical skills* refer to an ability to search for and process information. See Section 2.2 for full definitions.

experiment among a large sample of German recruiters who hire apprentices through our access to the BIBB Qualification Panel (Friedrich et al., 2022), a nationally representative annual survey of German firms.

In our experiment, we presented 1,958 recruiters, primarily firm owners and HR managers, with six fictitious applicants who applied for an apprenticeship position in the most common training occupation within the firm. We asked the recruiters to evaluate the probability of inviting these applicants for a job interview. During the experiment, we simultaneously randomised eight characteristics of the applicants, which included ethnicity (indicated by an ethnicity-specific surname), levels of social skills, knowledge about the profession, analytical and digital skills, gender, time of residence in Germany, and education level. This experimental setup created independent and exogenous variations in the applicants' characteristics, enabling us to causally identify recruiters' preferences for different ethnic groups and determine how the applicants' skills can mitigate potentially discriminatory behaviour of employers in the hiring process.

Our empirical approach is advantageous for our research question from at least three perspectives. First, our approach allows us to study the effect of the prospective policies regarding skill disclosure. The contemporary labour market places a significant emphasis on the importance of skills during the hiring process (Humburg & van der Velden, 2015). At the European level, several initiatives have been proposed to improve the disclosure of skills.² However, the effect of these initiatives and the optimal methods to reveal skills beyond conventional educational classifications remain open questions. Capturing the effects of skills disclosure through observational data or field experiments is challenging, given that these incentives are yet to be implemented in practice. Our empirical design, however, relies on hypothetical scenarios and enables us to examine the potential effects of such policies.

Second, administering our experiment as part of a larger survey granted us greater potential for heterogeneity analyses compared to correspondence studies in the field, where the characteristics of the recruiter and the firm often remain unobserved. Our

²See, e.g. European [Europass skill profile](#) and [Skill Passport](#).

survey includes information about the risk attitudes and sociodemographic characteristics of recruiters, as well as detailed data on firm size, industry, occupations, salary scale, and workforce composition. Although heterogeneity analyses offer descriptive evidence and should be interpreted with caution, they enable us to uncover the interplay between diverse factors that affect discrimination, producing insights that can guide more targeted and effective policies.

Third, our experiment offers full control over the information available to recruiters. One of the inherent shortcomings of the correspondence studies that are prevalent in the literature is the lack of control over certain relevant variables, particularly the number and the characteristics of competitors, which can result in biased estimates (Neumark, 2018; Phillips, 2019).³ By contrast, in our experiment we observed all alternatives that recruiters face so that we could account for them, thereby enhancing the precision and validity of our findings.⁴

In our study, we focus on the apprenticeship market, which is a major entry-level labour market for young people in Germany (Kubler et al., 2018). Apprenticeships are part of Germany's dual system of vocational training, where young people attend a vocational school for one to two days a week while gaining practical work experience in a firm. Completing an apprenticeship is essential for securing regular employment, especially for individuals without a higher education degree. Many jobs require having completed an apprenticeship.

Despite an educational component that differentiates apprenticeships from regular employment, the hiring procedures for apprentices largely resemble those for other employees. Applicants have to apply directly to the firm and employers are free to set their own selection criteria. Similarly to the regular job market, there is a large variation in the length and professionalisation of the recruitment process, and salaries vary significantly between occupations and industries. Given that firms have the discretion to decide

³In correspondence studies researchers send out fictitious resumes that are identical except for one variable of interest (e.g. the applicant's name) to various job openings and measure the call-back rates. See Neumark (2018) for an overview of these studies.

⁴While factorial survey experiments provide valuable insights into how recruiters form judgements, this approach is not without its limitations. In Section 5, we discuss the limitations of our study and provide an account of the strategies we employed to address them.

whether or not to offer apprenticeships, the apprenticeship market is highly competitive every year and many applicants are unable to secure an apprenticeship position (BIBB, 2021).

We chose apprenticeships for two major reasons. First, the inequality in training opportunities is especially disadvantageous for young workers because of its impairing effect on professional skills development and future career trajectories (Altonji et al., 2016; Arellano-Bover, forthcoming). More than two-thirds of the apprentices remain with the firm that trained them (BIBB, 2021), which makes pursuing an apprenticeship a significant stepping stone in finding the first job and shaping a career path. Hence, discrimination in the apprenticeship market can have a persistent distorting effect on future labour outcomes for young people.

Second, more than 60% of labour shortages in the EU are concentrated in VET occupations (McGrath, 2021), i.e. medium-level qualified occupations delivered by the traditional apprenticeship system. While migration is considered one of the solutions to the issue of labour shortages, for many young immigrants and immigrants with lower education, entering the apprenticeship system is often the only way to acquire a local qualification and find regular employment. Discrimination in the apprenticeship market can impede the labour market entrance of migrants and hinder their integration.

This study reports four main findings. First, we found evidence of discrimination against immigrant-origin minorities in the apprenticeship market. Applicants with ethnicity-specific surnames are on average 2.3 percentage points less likely to be invited to a job interview compared to applicants with German surnames. Second, we show that social skills moderate differences in evaluations between German and immigrant-origin minority applicants. This indicates that social skills suppress cultural stereotypes and signal cultural similarity that helps to alter discriminatory attitudes. Third, professional skills intensify differences between German and immigrant-origin minority applicants. There are larger disparities in job interview probabilities between German and non-German applicants at the top of the professional skills distribution, which may contribute towards poorer applicant-to-job match for young job applicants with a migrant background.

Fourth, we show that ethnic discrimination and the extent to which particular skills can moderate or intensify disparate treatment vary between ethnic groups.

We contribute to the existing literature in three ways. First, we add to the literature investigating why job seekers with a migrant background are discriminated against in the labour market and what factors can reduce ethnic discrimination. A theory of statistical discrimination suggests that a significant part of ethnic disparities is due to a lack of relevant productivity-related information in job applications and thus greater uncertainty about minority applicants (Arrow, 1973; Phelps, 1972). Although adding more detailed personal information is often expected to reduce uncertainties and discrimination, existing empirical evidence is very limited and has produced mixed results (Agerström et al., 2012; Baert et al., 2017; Dahl, 2022; Kaas & Manger, 2012; Oreopoulos, 2011; Thijssen et al., 2021).⁵ In this study, we contribute to this literature by investigating whether information about the applicant's cognitive and non-cognitive skills helps alleviate the gap in hiring probabilities between migrant and native job applicants.

Second, unlike many other experiments on ethnic discrimination (e.g., Bertrand & Mullainathan, 2004; Kaas & Manger, 2012; Protsch & Solga, 2017)⁶, we included applicants from several ethnic groups both within and outside of Europe in our experiment. This allowed us to investigate whether and how attitudes of recruiters differ towards various minority groups and what implications this may have for people with both EU and non-EU migration backgrounds. The literature addressing differences in discrimination patterns across ethnicities remains scarce (Lippens et al., 2023; Thijssen et al., 2021). Typically, field experiments investigate ethnic discrimination for job applicants from only one ethnic minority group, i.e. the most stigmatised. Growing within-EU and international migration have significantly increased diversity among immigrant populations and raised the relevance of exploring employers' attitudes toward various ethnic groups, both to gain a better understanding of the drivers of discrimination in hiring and to develop appropriate policy responses.

Third, our study is an improvement over existing experiments in terms of scale and

⁵See Table S1 in the Supplementary Materials for an overview of their findings.

⁶See Neumark (2018) for an overview of this literature.

external validity. In our experiment, a total of 11,748 evaluations were collected from a nationally-representative sample of actual hiring professionals and firm owners from different industries. The selection of occupations in our experiment reflects the total population of jobs in the German apprenticeship market. This gives us an advantage over correspondence studies, which usually limit themselves to a limited number of occupations. This, in turn, hinders the generalisability of their findings to a broader context (Neumark & Rich, 2019). This also grants us an advantage over existing vignette experiments, where the number of the included occupations was also fairly small (Lippens et al., 2023). Finally, the use of actual recruiters as respondents enhances the validity of our study, making it more representative of real-world hiring decisions, whereas Baert and De Pauw (2014), Blommaert et al. (2014), and Lahey and Oxley (2021) used students as respondents.

The remainder of this study is organised as follows. In Section 2, we explain the setup of the experiment and describe our data. Section 3 outlines the empirical strategy, Section 4 presents the results, and Section 5 concludes the paper.

2 Data

2.1 Experiment setup

To investigate how the interplay between applicants' skills and ethnicity affects hiring decisions, we conducted a factorial survey experiment. The experiment was embedded in the 2021 wave of the BIBB Establishment Panel on Qualification and Competence Development (short: BIBB Qualification Panel).⁷ The BIBB Qualification Panel is an annual survey of firms conducted by the German Federal Institute for Vocational Education and Training (BIBB). The survey collects data on the structure of the establishment and the

⁷Due to safety measures during the pandemic, respondents' answers were collected via computer-assisted telephone interviewing (CATI). A further motive to abandon the usual on-site interviewing practice was steering clear of possible selection into the sample since refusal to participate in a vis-à-vis interview in (post)pandemic time could be correlated with respondents' characteristics associated with discriminatory behaviour (e.g. risk attitudes). Although the use of personal interviews is more common in the literature, several studies also show the suitability of CATI for factorial experiments (Andernach & Schunck, 2014; Damelang & Stumpf, 2020).

characteristics of business administration and provides information on training activities in the establishments.⁸ The survey population is a nationally representative sample of German firms obtained from the database of the German Federal Employment Agency, which covers all establishments where at least one employee is subject to mandatory social insurance contributions in Germany.

Please imagine that your firm has an urgent need to fill in an apprenticeship position in occupation [Insert: relevant for the firm occupation (based on pre-survey)].

I will describe six applicants for this position. The applicants differ only in attributes mentioned in the descriptions and are similar across any other characteristics. For each applicant please indicate how likely this person will get an invitation to a job interview.

Applicant 1:

The person's surname is Weber, it is a man, who lives in Germany since the start of secondary school, has finished General secondary education, and has very good knowledge about profession, very good digital skills, very good analytical skills, and satisfactory social skills.

On a scale from 1 to 10 what is the probability that this person will get an invitation to a job interview in your firm?

1 stands for „very unlikely“, 10 stands for „very likely“.

very unlikely	very likely
<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/> <i>don't know</i>	

Fig. 1. Example of the experiment task.

Notes: English translation. See section C in the online Supplementary Materials for German text.

To avoid unrealistic situations in our experiment, we included only recruiters in firms that offer apprenticeship positions. The respondents of the survey, mainly firm owners and HR managers, were asked to evaluate six fictitious applicants sequentially as if they had applied for an apprenticeship position in their firm. To increase the relevance of this task for recruiters, we used the information about their firm gathered in the pre-survey and modified the applicants' descriptions to match the specific occupation for which the

⁸More information about BIBB Qualification Panel is available under <https://www.bibb.de/en/1482.php>

firm provided apprenticeship training accurately. We instructed our respondents that the applicants differed only in characteristics presented in the descriptions, and did not differ with respect to any other features, including fluency in the German language and current place of residence. For each applicant, respondents rated the probability of invitation to the next stage of the recruitment process (job interview) on a scale from 1 (very unlikely) to 10 (very likely). Figure 1 provides an example of the task that was presented to respondents. After evaluating six profiles, the respondents filled in a short post-experimental survey in which they were questioned about their personal characteristics, such as position held in the firm, age, gender, country of birth, and risk attitude.

The external validity of our study benefits from the fact that our sample of respondents represents the target population of hiring professionals who were asked to perform a task that resembles their usual work (Hainmueller et al., 2015). To avoid a cognitive burden, we kept the number of characteristics presented in applicants' descriptions and the number of questions per respondent sufficiently small (Auspurg & Hinz, 2014). We also provided our respondents with the handout materials ahead of the interview (see the experiment introduction text and handout materials in Section B of the online Supplementary Materials).

2.2 Descriptions of applicants

To mimic real-life hiring decisions as closely as possible, we made the applicants different in terms of eight characteristics that were shown by the existing literature to be relevant during a hiring process(Damelang & Stumpf, 2020; Humburg & van der Velden, 2015; Mergener & Maier, 2019). They included an indication of the applicants' ethnicity, social skills, knowledge about the profession, digital and analytical skills, applicants' gender, time of residence in Germany, and education level. Table 1 presents an overview of all characteristics and their values.

Table 1
Applicant characteristics.

Factors	Dimensions	Frequencies
Ethnicity (signalled by surname)	1. German (Weber, Schneider) 2. Dutch (de Vries, van den Berg) 3. Arab (Al Numan, Khaled) 4. Polish (Jankowski, Kowalski)	24.84 25.22 25.14 24.80
Social skills	1. Satisfactory 2. Good 3. Very Good	33.13 33.43 33.44
Knowledge about the profession	1. Satisfactory 2. Good 3. Very Good	33.55 33.18 33.27
Digital skills	1. Satisfactory 2. Good 3. Very Good	33.88 33.04 33.09
Analytical skills	1. Satisfactory 2. Good 3. Very Good	33.62 33.47 32.91
Gender	1. Male 2. Female	50.24 49.76
Time of residence in Germany	1. Since birth 2. Since secondary school 3. Since recently	33.40 33.30 33.30
Education level	1. Lower secondary 2. General secondary 3. Higher secondary 4. Unfinished higher education	25.12 24.78 24.90 25.20

Notes: The table lists characteristics included in descriptions of the fictitious applicants along with the corresponding values and their frequencies in the final sample. *Lower, general* and *higher secondary* education levels correspond to German *Hauptschule*, *Realschule* and *Abitur*, respectively.

Rather than mentioning their ethnicity directly, we signalled the applicants' ethnicity by using ethnicity-specific surnames. We did not use first names in our experiment, since recent research shows that they can also signal characteristics other than ethnicity, such as perceived social class or the migration status of the individual (Baert et al., 2022; Van Borm et al., 2022). For the sake of plausibility for the recruiters, we selected ethnicities that come from neighbouring countries and/or represent the largest recent migrant groups. To reflect the growing diversity among the immigrant population, we included applicants of Western European (Dutch), Eastern European (Polish), and non-

European (Arab) ethnicity, which vary with respect to cultural proximity to the German population.

We randomly assigned surnames that indicate German, Dutch, Polish, or Arab ethnicity to the applicants' descriptions. The surnames we used belong to the top ten most widespread surnames among the corresponding nationalities. The data on the frequency and ethnic origin of surnames come from the database of Forebears.io, and are based on information regarding the country where people with a particular surname currently reside.⁹ To avoid undesirable positive or negative connotations, we did not use surnames that might be associated with famous persons, for example widely known Dutch or German soccer players, politicians, or religious leaders. We also excluded surnames that proved confusing to respondents during the pilot study for our experiment.¹⁰

The descriptions of the applicants also contained information about cognitive and noncognitive skills, shown to be of significant importance for the hiring process in general and the hiring of apprentices in particular (Heckman et al., 2006; Protsch & Solga, 2015). Based on the literature, we distinguished between social skills (Borghans et al., 2014; Deming, 2017; Homburg & van der Velden, 2015; Piopiunik et al., 2020; Protsch & Solga, 2015), knowledge about the profession (Biesma et al., 2007; Homburg & van der Velden, 2015; Nunley et al., 2016; Tobback et al., 2020), digital skills (Piopiunik et al., 2020) and analytical skills (Biesma et al., 2007; Piopiunik et al., 2020; Protsch & Solga, 2015) in our experiment.

We defined the applicant's social skills as an ability to work in teams, communicate and cooperate effectively, and cultural awareness in general. Potential employers could infer the applicants' social skills from extracurricular activities listed in the CV, such as team sports and volunteering experience, or from school grades for communication skills and teamwork (in German *Kopfnoten*). These school grades are included in the transcript of records in some German regions (Protsch & Solga, 2015). Knowledge about the

⁹URL: <https://forebears.io/onograph/documentation>, accessed on 7 August 2020

¹⁰To ensure that our experiment was clear and well-constructed, we ran pilot study with 30 German employers. Following the employers' feedback during this pilot study, we refined the list of ethnicity-specific last names for our experiment by removing those that sounded confusing to participants. The results of this pilot study are available upon request.

profession reflects applicants' awareness of job requirements and job-related experience, which employers could infer from internships, side jobs, professional orientation courses or related extracurricular activities listed in the application documents. Digital skills reflect the use of relevant software, computer and programming skills, as signalled by e.g. IT competences reported in the CV and grades for the Informatics school course. Analytical skills imply reflectiveness, an ability to search for and process information, as well as an ability to approach problems from different angles. The analytical skills of the applicant could be inferred from academic performance, hobbies and extracurricular activities, such as playing chess.

To facilitate the comprehension of the quality of the skills by the respondents, all skills were measured on a German school grade scale, using the values 'very good', 'good' and 'satisfactory'. We introduced these grades to recruiters as if they were their own judgement about applicants' skills based on application documents. All recruiters were provided with a visual aid containing descriptions of the skills ahead of the interview (see the handout text in Section B of the online Supplementary Materials).

Further characteristics included gender, time of residence in Germany and education level of the applicant. The gender of the applicant was either female or male. Time of residence in Germany took the values "since recently" signalling for a newly arrived first-generation migrant, "since the start of secondary school" for a first-generation migrant who nevertheless partly followed German educational curriculum, and "since birth". The education level of the applicant corresponded either to lower secondary (German *Hauptschule*), general secondary (*Realschule*), higher secondary (*Abitur*), or unfinished higher education. While the first three levels of education are commonly used in the existing literature on apprenticeships, we also introduced the value of "unfinished higher education" to reflect the rise in higher education dropout rates in Germany, as documented by several studies (Heublein et al., 2017; Heublein, 2014), which comprises about 29% of the enrolled students. Applicants who discontinue their higher education often need to complete vocational training to enter the labour market and therefore have to compete with school graduates for apprenticeship positions. However, to the best of our

knowledge, none of the existing studies looked at the performance of this group in the apprenticeship market.

2.3 D-efficient design

Since the number of all possible combinations of values of applicant characteristics that can be included in the description was large ($4 \times 2 \times 3 \times 4 \times 3 \times 3 \times 3 \times 3 \times 3 = 7776$ unique combinations) and exceeded the expected number of respondents, we implemented a D-efficient design using the algorithm described in Auspurg and Hinz (2014). This algorithm selects combinations of characteristics' values with the most statistical power and guarantees that they are nearly orthogonal (i.e. are mutually uncorrelated and appear in the survey equally often).¹¹ A D-efficient design enhances statistical precision, thus allowing us to use fewer combinations of characteristics and fewer judgements from our respondents to attain the same amount of statistical power.

Using this algorithm, we selected 750 unique combinations of characteristics' values with the highest D-efficiency score.¹² We chose our design in a way that allowed us to estimate all main effects and all possible two-way interactions of the applicant's characteristics, as well as selected three-way interactions.¹³ We then divided these combinations into 125 blocks that were randomly assigned to respondents, with each block consisting of six unique descriptions of the fictitious applicants. To account for possible order and fatigue effects, we randomly varied the order in which the descriptions were shown and the order of characteristics within descriptions across respondents.

¹¹See column 3 of Table 1 for the frequencies of characteristics' values in our survey. All dimensions appear in a survey with similar frequencies, which proves the successful randomisation and balance of the experiment.

¹²The D-efficiency score of our design equaled 99.3, while the recommended level is above 90 (Auspurg & Hinz, 2014)

¹³Our design allowed us to estimate any tree-way interactions that included at most one skill characteristic, e.g. the interaction term *social skills* \times *gender* \times *ethnicity* could be estimated, while the term *social skills* \times *analytical skills* \times *ethnicity* could not.

2.4 Descriptive statistics

In total, 2,120 recruiters took part in our experiment. Among them, 90 respondents made less than six evaluations. We removed observations from respondents who had not answered all six questions or who had filled in the survey in an implausibly short amount of time.¹⁴ In total, we removed answers from 162 respondents from our sample.¹⁵ Thus, our final sample consisted of 11,748 evaluations of fictitious apprenticeship applicants made by 1,958 recruiters.

The distribution of the probability of interview, our main outcome variable, tended to be skewed to the right with a mean equal to 6.8 on a 10-point Likert scale (see Figure S1 in the online Supplementary Materials). The evaluations of German and Non-German applicants had a similar distribution, however, non-German applicants had a higher probability of getting below-average evaluations compared to German applicants.

In Table A1, we present descriptive statistics for the characteristics of the recruiters (A), the characteristics of the firms (B), and the evaluation of the probability of an interview (C). In the first column, we present the descriptive statistics for the entire sample. In order to check for successful randomisation, we also present the descriptive statistics for the two subsamples with German (column 2) and non-German (column 3) surnames of fictitious applicants.

As shown in column 1, the average age of the recruiters in our sample was 47 years and 44.9% of our respondents were women. About 19.8% of firms in our sample were located in East Germany, and 15.1% of the firms reported having employees with a migrant background. Due to successful randomisation, there was no difference in the characteristics of recruiters and firms between sub-samples of vignettes with German and non-German surnames (see column 4 of Table A1 for a formal test). This means that our fictitious apprenticeship applicants with German and non-German surnames were evaluated by recruiters who were similar in terms of age and gender, and who worked in firms of comparable size, location, industry and workforce composition. Column 4 also presents descriptive evidence for the difference in evaluations between applicants with

¹⁴As suggested by Leys et al. (2013), we used < 2.5 Median Absolute Deviations (MAD) as a threshold.

¹⁵Including these observations does not change the results of our analysis.

ethnicity-specific and German surnames. The result suggests that applicants with non-German surnames were rated 0.15 points lower on average than their competitors with German surnames, which corresponds to a lower probability of 2.2 percentage points of being invited to a job interview compared to the average German applicant.

3 Estimation strategy

To estimate the results, we employed a random effects model, considering the dependencies of answers at the respondent level.¹⁶ We clustered the standard errors at the respondent level and controlled for the position in which each question was shown in order to account for potential order effects. Our baseline model takes the following form:

$$(1) \quad Y_{ir} = \alpha_0 + \alpha_1 E_{ir} + \alpha'_2 S_{ir} + \alpha'_3 C_{ir} + \alpha'_4 X_r + \mu_{ir} + \varepsilon_{ir}$$

where Y_{ir} is an evaluation of the probability to invite applicant i for a job interview given by recruiter r on a 10-point Likert scale; E_{ir} denotes the applicant's ethnicity signalled by surname; S_{ir} is a vector of the applicant's skills; C_{ir} is a vector of other applicant characteristics shown to each recruiter, namely gender, time of residence and education level; X_r represents a vector of the recruiter and firm level controls that include industry, firm size, location (East or West Germany), and gender and age of the recruiter; μ_{ir} stands for a position in which the applicant's description appeared during the interview; and ε_{ir} is an independent error term, clustered at the respondent level. Thus, the coefficients in the vectors α'_{1-3} capture the impact of a change in the characteristic of the respective applicant on the probability of being invited to a job interview.

To investigate how applicant skills moderate the level of ethnic discrimination in hiring, we estimated a model with interactions between the applicant's ethnicity signalled by surname and skills, accounting for all other applicant characteristics and selected firm,

¹⁶Random effects models are recommended for factorial survey experiments, for they are better suited for the hierarchical structure of data where each respondent evaluates several applications (Auspurg & Hinz, 2014). However, our results are robust to alternative estimation strategies and to an exclusion of the respondent- and firm-level controls. Table S2 in the online Supplementary Materials provides estimates of Equation 1 using different estimation methods.

respondent and order controls:

$$(2) \quad Y_{ir} = \beta_0 + \beta'_1(E_{ir} * S_{ir}) + \beta_2 E_{ir} + \beta'_3 S_{ir} + \beta'_4 C_{ir} + \beta'_5 X_r + \mu_{ir} + \eta_{ir}$$

where $E_{ir} * S_{ir}$ stands for an interaction term between the applicant's ethnicity E_{ir} and skill level S_{ir} , and the elements of the coefficient vector β'_1 are our parameters of interest. The interaction effects β'_1 capture the degree to which the quality of a particular skill is consequential for the relative difference in evaluations between minority and majority applicants. If it reduces discrimination, the estimates should be positive.

To benchmark the magnitude of the effects, we computed average marginal effects from the interacted model and reported the impact as a percentage change in the probability to get an invitation to a job interview relative to the level of German applicants.

4 Results

4.1 The role of the applicant's characteristics

First, we analysed the effect of the applicant's characteristics on the probability of invitation to a job interview and checked the prevalence of ethnic discrimination in the German apprenticeship market. Figure 2 depicts the effects estimates and confidence intervals from Equation 1.¹⁷ For convenience of interpretation, we reported the impact both in Likert points (as in tables and figures) and as a percentage change in the probability of an invitation to a job interview relative to the reference level of the corresponding characteristic.

We found evidence of ethnic discrimination in the German apprenticeship market. Our findings suggest that a relatively small, yet statistically significant, part of the variation in evaluations can be attributed to the ethnic surname of the applicant. On average, applicants with a non-German surname received evaluations that were 0.16 Likert points lower evaluations, which translates into a 2.3 percentage points lower probability of re-

¹⁷We report the corresponding quantitative estimates in Table S3 in the Appendix.

ceiving an invitation to a job interview compared to an average German applicant.

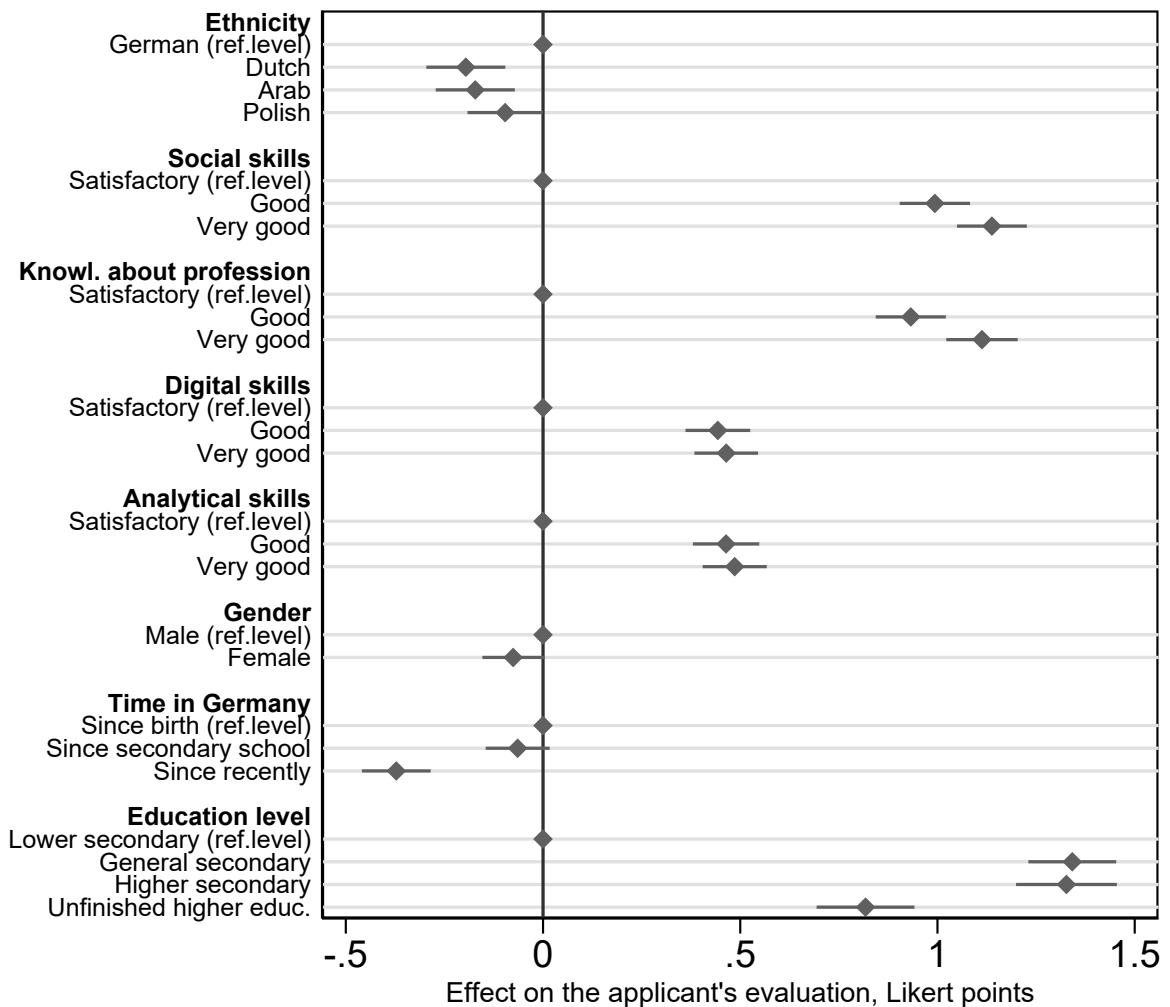


Fig. 2. Impact of the applicant's characteristics on access to apprenticeships.

Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure presents the effects of the applicant's characteristics on the probability of receiving an invitation to a job interview obtained from Equation 1. The effects were measured on a 10-point Likert scale. Standard errors were clustered at the respondent level and 95% confidence intervals are shown on the graphs.

The ethnic penalty varied depending on the signalled ethnicity. We found that applicants with Dutch and Arab surnames faces a 2.8 and 2.5 percentage points (0.20 and 0.17 Likert points) lower probability of an invitation to a job interview, respectively, while for applicants with Polish surnames the probability of an invitation was only 1.4 percentage points (0.10 Likert points) lower.¹⁸ These differences are in line with the ethnic hierarchy theory (Sidanius & Pratto, 1999) proposed in social psychology, which

¹⁸p-value for the t-test for the difference in the effect between Polish and Dutch applicants equals 0.045, and 0.125 for the difference between Polish and Arab applicants.

assumes that a hierarchy of preferences towards particular ethnic groups exists in society. However, our findings cannot support the hypothesis in this literature that the position of ethnic minorities depends on their socioeconomic status and the degree to which they culturally deviate from the native-majority population (Hagendoorn, 1995), for it appears that the Western minority group represented by applicants with Dutch surnames in our study faced similar rates of discrimination as more culturally remote applicants with Arab surnames.

Skills played an important role in the recruiters' decision-making process. The results suggest that the recruiters' preferences were strongest for applicants with higher levels of social skills and knowledge about the profession. Applicants with an above-satisfactory level of social skills and knowledge about the profession received evaluations that were 18.2 and 16.4 percentage points (1.1 and 1.0 Likert points) higher on average than the evaluations of applicants with a satisfactory level of these skills. Possessing an above-satisfactory level of digital and analytical skills raised the applicant's evaluation by 6.9 and by 7.3 percentage points (0.45 and 0.47 Likert points), respectively. There was only a small extra benefit for having a very good level of skills compared to a good level.

The gender of the applicant had the smallest influence on the probability of an invitation to a job interview: being female reduced the applicant's evaluation by 1.2 percentage points (0.08 Likert points) compared to the evaluations male applicants, with the result being only marginally significant. Applicants who had recently arrived in Germany received evaluations that were 5.7 percentage points (0.37 Likert points) lower than those who had resided in Germany since birth. However, there was no significant difference in the probability of an invitation to a job interview between applicants who had resided in Germany since the beginning of secondary school and since birth. Among the applicant's characteristics, the education level had the strongest impact on the probability of an invitation to a job interview. Applicants with an above-lower secondary education received evaluations that were 20.4 percentage points (1.2 Likert points) higher on average. However, applicants with an unfinished university degree were rated significantly lower compared to general and higher secondary school graduates, suggesting that recruiters

perceived interruption of higher education studies as a negative signal.

4.2 Can skills alleviate ethnic discrimination?

We went on to explore whether the effect of ethnicity could be moderated by the quality of the applicant's skills. Table 2 presents the results of regression analysis with interaction effects between skills and ethnicity described by Equation 2.

Table 2
Interaction effects between skills and ethnicity.

	(1)	(2)
Social skills \times non-German	0.103** (0.052)	
Very good \times non-German		0.204** (0.100)
Good \times non-German		0.151 (0.107)
Knowledge about the profession \times non-German	-0.108** (0.054)	
Very good \times non-German		-0.215** (0.107)
Good \times non-German		-0.113 (0.111)
Digital skills \times non-German	-0.045 (0.052)	
Very good \times non-German		-0.089 (0.104)
Good \times non-German		-0.047 (0.102)
Analytical skills \times non-German	0.041 (0.050)	
Very good \times non-German		0.085 (0.101)
Good \times non-German		-0.082 (0.098)
Non-German	-0.142 (0.211)	-0.109 (0.126)
<i>(Satisfactory \times non-German - omitted)</i>		
Observations	11,748	11,748

Notes: The table presents coefficients β'_1 for interaction terms between skills and non-German ethnicity from the fully interacted model described by Equation 2. Regressions included all applicant characteristics, firm size, industry, location, recruiter's age and gender, and question order dummies as controls (see Table S4 in the online Supplementary Materials for an extended version of this table). The model in column 1 assumes a linear effect of skill quality, while the model in column 2 presents interaction effects at each level of skill separately. Standard errors clustered at the recruiter level (1,958 clusters) are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A positive coefficient before the interaction term between social skills and non-German

ethnicity in Table 2 suggests that investment in higher levels of social skills helps to narrow the gap between German applicants and applicants with ethnicity-specific surnames. The estimates suggest that for every step of increase in the level of social skills, the gap in job interview probability between non-German and German applicants decreases by 1.4 percentage points (0.10 Likert points). Conversely, our findings show that the disparity between non-German and German applicants increases at higher levels of knowledge about the profession. A one-step increase in the level of knowledge about the profession widens the gap between non-German and German applicants with a similar level of skills by 1.6 percentage points (0.11 Likert points).

In Figure 3, we offer graphical support for these results and demonstrate the relative performance of non-German applicants compared to German applicants with identical personal characteristics. Figure 3 presents the average marginal effects of non-German ethnicity on job interview probability at each level of skills.

As Figure 3a suggests, minority applicants with a satisfactory level of social skills on average had a 3.8 percentage points (0.23 Likert points) lower probability of an invitation to a job interview compared to German applicants with the same level of skills. As the quality of social skills improved, the gap narrowed. Thus, non-German applicants with a good level of social skills had a 2.1 percentage points (0.15 Likert points) lower probability of an invitation compared to their German counterparts, while there was no statistically significant difference in evaluations between non-German and German applicants with a very good level of social skills. By contrast, the gap in evaluations between German and ethnic applicants increased as the level of knowledge about the profession grew (see Figure 3b). While the differences in evaluations at the satisfactory level of knowledge about the profession were small and not statistically significant, they equalled 2.5 percentage points (0.18 Likert points) at the good level and 4.0 percentage points (0.29 Likert points) at the very good level.

Digital skills also exhibited a negative pattern, but the size of the effect was small (see Figure 3c). The gap between minority and comparable German applicants widened as the level of digital skills grew, equalling 1.8 percentage points (0.11 Likert points)

at the satisfactory level of digital skills, 2.2 percentage points (0.16 Likert points) at the good level, and 2.7 percentage points (0.20 Likert points) at the very good level. For analytical skills, the difference in probability of an invitation between non-German and German applicants increased at the good level of skills, equalling 3.3 percentage points (0.24 Likert points) compared to 2.5 percentage points (0.15 Likert points) at the satisfactory level, but disappears at the very good level (see Figure 3d).¹⁹

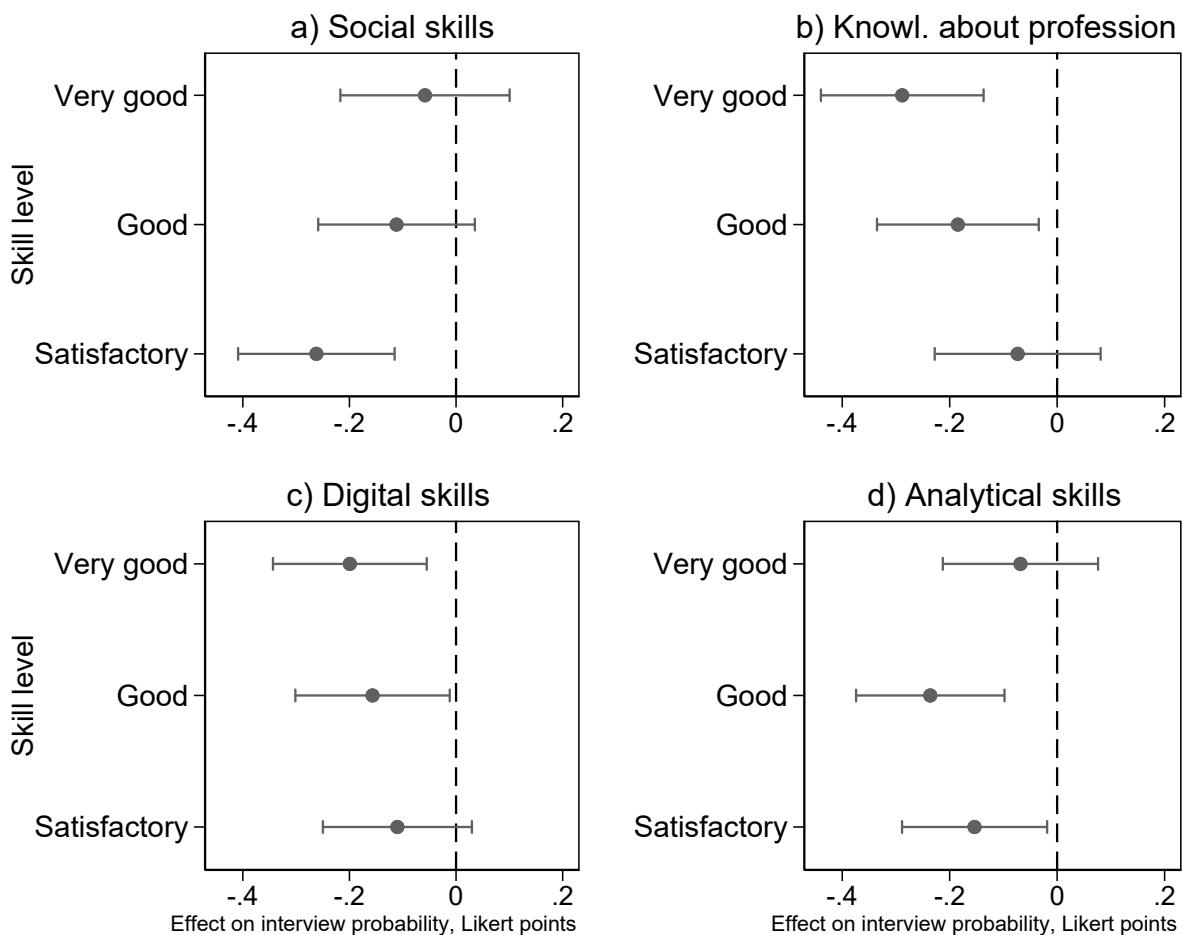


Fig. 3. Interaction between skills and non-German ethnicity, marginal effects.

Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure plots marginal effects for non-German applicants at different levels of skills and knowledge about the profession relative to comparable applicants with German surnames. Estimates were obtained by calculating marginal effects from a regression described by Equation 2. Standard errors are clustered at the respondent level, and 95% confidence intervals are shown on the graphs.

In total, our results show a heterogeneous effect of skills on ethnic discrimination.

¹⁹The positive effect at the highest level of analytical skills comes from a lower probability of an invitation to a job interview for German applicants, while the evaluation of non-German applicants remained unchanged (see Figure S2 in the online Supplementary Materials). One possible explanation for this result is that employers may perceive German applicants with a very good level of analytical skills as overqualified for an apprenticeship position.

On the one hand, the evidence supports the notion that the quality of social skills has a positive effect on the relative evaluation of minority applicants and helps alleviate ethnic discrimination. On the other hand, our results also suggest that disparate treatment intensifies at higher levels of knowledge about the profession and digital skills. This suggests that immigrant-origin minorities can be highly qualified, yet face significant discrimination, which might contribute toward less efficient applicant-to-job matches for minority applicants.

4.3 Effect heterogeneity

4.3.1 *Differences across ethnicities*

We then checked whether the effect of the applicant’s skills varied across the signalled ethnicities. The results presented in Figure A1 show that the extent to which a particular skill can moderate or intensify discrimination varies considerably between ethnic groups.

The moderating effect of social skills was most pronounced for applicants of Polish and Dutch descent. For applicants with Polish and Dutch surnames, improving the quality of social skills helped to narrow the gap in interview probability between them and comparable German applicants considerably. However, we found no evidence of a moderating effect of social skills for applicants with Arab surnames: for them, acquiring a better level of social skills did not reduce discrimination. One possible explanation for the absence of any positive effect for applicants of Arab descent is that they face more taste than statistical discrimination. Taste discrimination is not driven by the recruiter’s uncertainty about relevant productivity-related characteristics and therefore cannot be resolved by giving productivity-related information (Guryan & Charles, 2013). Consequently, our findings suggest that although social skills help to reduce the level of ethnic discrimination, they are not equally effective across different ethnic groups.

The discrimination-intensifying effect of knowledge about the profession prevailed among applicants with Dutch and Arab surnames. Recruiters gave applicants with Dutch and Arab surnames less credit for the better quality of knowledge about the profession compared to German applicants, leading to larger relative differences in interview prob-

ability at the top of the skills distribution. We did not observe similar negative patterns for Polish applicants. To explain this surprising negative effect, we estimated differences in marginal returns to knowledge about the profession depending on the applicant's time of residence in Germany. Due to cross-country differences in labour markets and labour market culture, employers might place less value on professional competencies acquired abroad compared to those acquired locally, which could explain our findings.

The results presented in Figure A2 support this hypothesis and show that marginal returns to knowledge about the profession are positively related to the applicant's time of residence. We observed no statistically significant differences in evaluations and marginal returns to knowledge about the profession for applicants who had resided in Germany since birth. However, for applicants who had not resided in Germany since birth, the relative differences in the probability of an invitation to a job interview increased as the quality of knowledge about the profession grew. Thus, those who had received professional experience abroad were worse off compared to local applicants because their knowledge about the profession was discounted. The fact the negative interaction between ethnicity and knowledge about the profession was evident for applicants with Dutch and Arab surnames, but not for those with Polish surnames, further suggests that employers value knowledge about the profession differently, depending on the country where it was acquired. We did not find significant interactions with time of residence for social, analytical and digital skills, suggesting that, unlike knowledge about the profession, employers see these types of skills as more transferable across national contexts.

4.3.2 *Heterogeneity by firm, recruiter and occupation characteristics*

Last, we investigated whether certain respondent, firm, or occupation characteristics affected the level of ethnic discrimination in the apprenticeship market and influenced the evaluation of non-German applicants' skills.²⁰

Table A2 reports the results of the multivariate regression analysis, which included

²⁰Note that interactions concerning occupation, firm and respondent characteristics cannot be given a causal interpretation, as they were not part of the randomisation process in the experiment and may therefore correlate with other unobservable characteristics.

interaction terms between the applicant's ethnicity and firm, recruiter, and occupation characteristics. In column 1, we estimated a baseline model where we regressed our outcome variable (interview probability on a scale of 1-10) on all applicant characteristics and a set of firm, recruiter, and occupation characteristics without any interaction terms. Consequently, in columns 2-4 we added different sets of interaction terms to our model. In column 2, we included interactions between the applicant's ethnicity and firm characteristics. We found a marginally significant negative interaction effect between the applicant's ethnicity and the size of the firm. This result implies that ethnic discrimination is more pronounced in small (<50 employees) enterprises, which is in line with the existing evidence (Baert et al., 2017; Carlsson & Rooth, 2007; Kaas & Manger, 2012). We did not observe any significant mediating effects of the firm's location and the share of employees with migration backgrounds in the firm. As the results in column 3 suggest, we also did not find any significant interaction effects between the applicant's ethnicity and the recruiter's age, gender, or risk attitude.

In column 4, we present the interaction effects between the applicant's ethnicity and occupational characteristics, such as a measure of occupational status, the Standard International Occupational Prestige Scale index (SIOPS), and whether the occupation requires communication with clients. The SIOPS scale is an internationally standardised measure of occupational status generated by averaging the prestige scores from national and local surveys conducted in 60 countries, where respondents were asked to rate occupations with respect to their social prestige (Ganzeboom & Treiman, 2003).²¹ The results suggest that ethnic discrimination is more prevalent in medium-status occupations, compared to low- and high-status ones.²² We did not find significant differences between occupations that did and did not require communication with clients. Finally, column 5 presents results

²¹In our sample, house and textile cleaners had the lowest SIOPS score of 25.2, while the highest score of 62.6 was attributed to programmers and legal professionals. The average score was 45.1 and applied, for example, to mechanics and booksellers.

²²The results are similar if we use the International Socio-Economic Index of occupational status (ISEI) as an alternative measure of occupational status (see Figure S3 in the online Supplementary Materials). The ISEI captures the socioeconomic status of occupations based on survey information about the earnings and education of individuals in particular occupations. To merge SIOPS and ISEI indexes with BIBB Qualification Panel data, we relied on the SIOPS - ISEI - ISCO-08 conversion provided by Ganzeboom (2010) and the KldB 2010 - ISCO-08 conversion key provided by the German Federal Employment Agency.

from the model with a robustness check where all interactions were included together.

The results are similar to those from previous models.

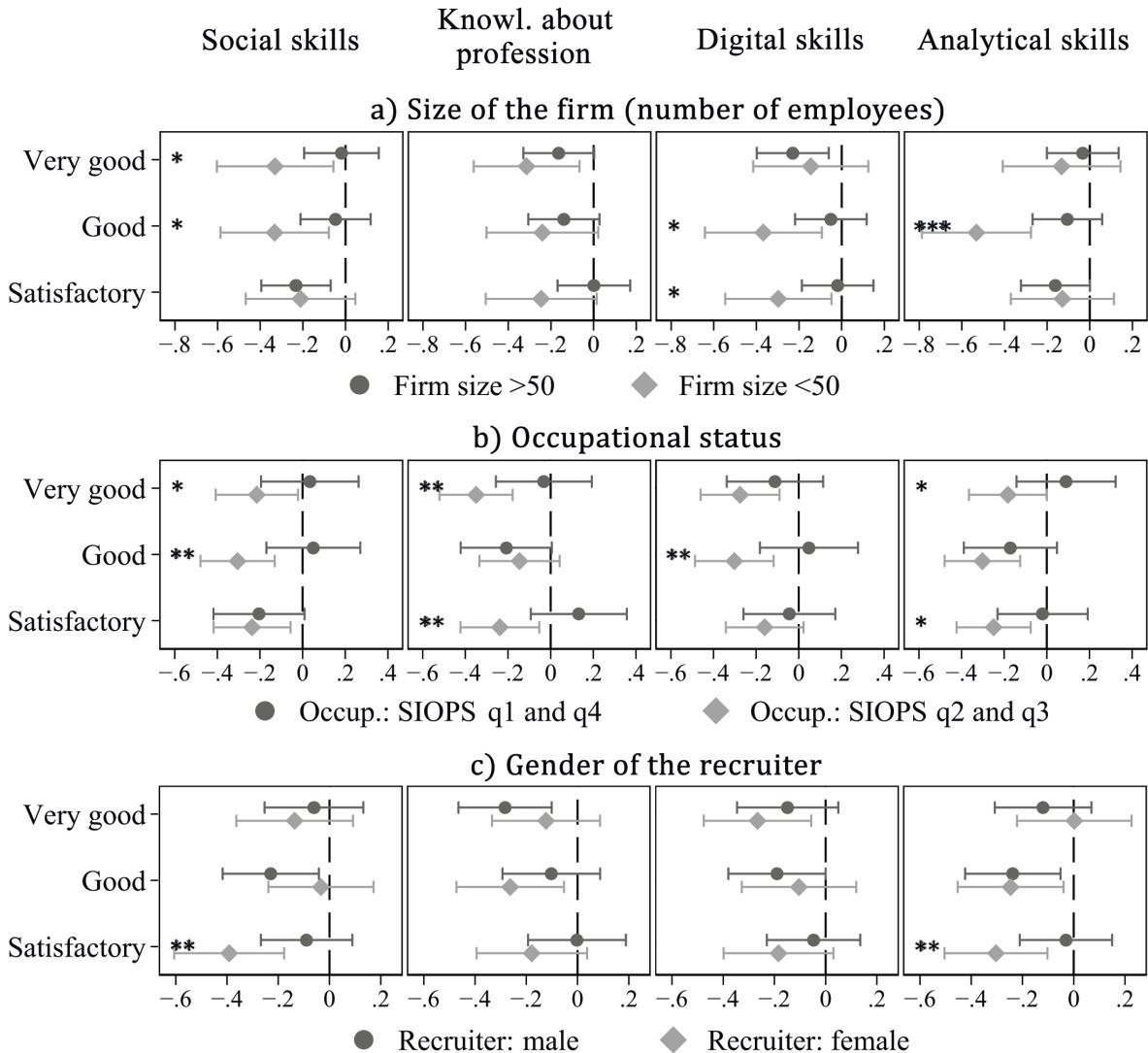


Fig. 4. Differences in evaluations of non-German applicants' skills by firm, recruiter and occupational characteristics.

Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure shows marginal effects for non-German applicants at different levels of skills and knowledge about the profession relative to comparable applicants with German surnames. The effects on the x-axes were measured on a 10-point Likert scale. Estimates were obtained by calculating marginal effects from a regression with three-way interactions between skills, ethnicity, and firm, recruiter, and occupational characteristics. Standard errors are clustered at the respondent level and 95% confidence intervals are shown on the graphs. The stars on the left side of the graphs show the level of significance of the difference in estimates by the corresponding firm, recruiter and occupational characteristic: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To demonstrate how the evaluation of skills of non-German applicants varies with respect to firm, recruiter, and occupation characteristics, we calculated marginal effects

using a regression model incorporating three-way interactions. Figure 4 presents the results of this analysis for the size of the firm (graph (a)), the occupational status (graph (b)) and the gender of the recruiter (graph (c)).

Our findings indicate that recruiters in small firms (<50 employees) assign lower evaluations to non-German applicants with lower levels of digital skills compared to recruiters in larger firms. Additionally, recruiters in small firms did not award extra credit for higher-quality social skills among non-German applicants. While improved social skills helped narrow the gap in probability of invitation to a job interview between German and non-German applicants for larger firms, this gap remained unchanged for small firms.

Social skills also did not help to close the gap between German and non-German applicants in middle-status occupations when compared to low- and high-status occupations. Furthermore, for middle-status occupations, the gap in the probability of invitation to a job interview between German and non-German applicants was more pronounced at a satisfactory and at a very good level of knowledge about the profession and analytical skills. The non-German applicants were also rated lower than their German counterparts at better levels of digital skills. However, for digital skills, the differences compared to low- and high-status occupations were only marginally significant.

Hence, the greater prevalence of ethnic discrimination within small firms and in middle-status occupations can be linked in part to differences in how the skills of non-German applicants are assessed. One aspect contributing to the higher ethnic disparities was the difference in the assessment of social skills. In bigger firms and in low- and high-status occupations, recruiters treated non-German applicants the same as German ones if they possessed at least a good level of social skills. In small firms and in middle-status occupations, improvement of social skills did not help to close the gap between non-German and German applicants.

Finally, our results indicate that female recruiters tend to assess non-German applicants with satisfactory levels of social and analytical skills more critically than their male counterparts. There were no differences between female and male recruiters when assessing applicants with a higher skill quality. We did not find significant variation in

the assessment of the skills of non-German applicants across other firm, recruiter and occupation characteristics included in Table A2. The results for these characteristics are presented in Figure S4 in the online Supplementary Materials.

5 Conclusion

Ethnic discrimination in hiring can be a barrier to the labour market integration of job seekers with migration backgrounds. It is particularly harmful to young job applicants whose impaired access to entry-level positions can reduce further career opportunities and affect labour market trajectories over the course of their lives. In this light, finding a way to reduce discrimination is of significant importance from a policymaking standpoint. Using data from a survey experiment among a nationally representative sample of German firms who hire apprentices, this study explores how the interplay between applicants' skills and ethnicity affects hiring decisions and what applicant skills can alleviate ethnic discrimination.

Our findings indicate that social skills are particularly effective in alleviating ethnic discrimination. Individuals' investment in social skills and incorporating social skills training as part of integration programmes can help reduce ethnic disparities in the probability of an invitation to a job interview. We further show that the extent to which social skills can moderate ethnic discrimination varies between ethnic groups. For applicants with Arab surnames, signalling a better quality of social skills is less effective in reducing ethnic discrimination, which arguably might point to higher levels of taste discrimination against Arab applicants. Lastly, we find evidence of skills discounting for first-generation migrant applicants' knowledge about the profession. The effect is most pronounced among applicants who have recently arrived in Germany and is not evident for those who have resided in Germany since birth. This suggests that recruiters tend to put less value on knowledge about the profession acquired abroad, leading to greater disparities between immigrant and native applicants at higher levels of skill quality.

The findings of our study offer valuable insights for policymakers and individuals

alike. First, our results indicate that interventions and training programmes focused on enhancing social skills can serve as an effective strategy to mitigate ethnic disparities in the apprenticeship market. This highlights the importance of addressing social skills development as a means to reduce discriminatory biases. However, it is important to note that our findings also suggest that there is no universal solution that applies to all ethnic groups. We discovered that, for certain ethnic groups, the potential to overcome discriminatory biases by signalling higher-quality social skills is limited. This underscores the need for targeted and tailored approaches that take into account the specific challenges faced by different ethnic groups.

Our study is not without limitations. The first challenge stems from the potential for hypothetical bias, as participants' responses in hypothetical scenarios may not accurately mirror their real-world behaviour and decisions. In our study, we mitigated this bias by employing two strategies. First, we conducted surveys with respondents in their natural working environments and limited our sample to real-life recruiters. That way, the tasks presented to respondents closely resembled their actual job responsibilities, enabling them to relate to the experimental scenarios better and respond more realistically. Second, we made the hypothetical scenarios sufficiently complex by simultaneously manipulating various applicant characteristics, in order to approximate the complexity of real hiring decisions. Previous research has shown that respondents' decisions in such experiment designs are highly correlated with their actual behaviour (Baert & De Pauw, 2014; Hainmueller et al., 2015).

Second, our experimental approach could be subject to a social desirability bias. Social desirability bias refers to the tendency of individuals to provide responses that they believe are socially acceptable or desirable, rather than reflecting their true beliefs or attitudes. To mitigate the influence of social desirability bias in our experiment, we presented participants in our study with a limited number of descriptions of fictitious applicants that varied over multiple significant characteristics, making it harder for participants to identify socially desirable answers (Auspurg & Hinz, 2014; Hainmueller et al., 2015). To check the extent to which our findings were affected by socially desirable answers, we

repeated our baseline analysis with a limited sample of responses that included only evaluations of the first fictitious applicant presented to each recruiter. The results of this exercise were largely in line with the results we found in the total sample, suggesting that social desirability bias played a lesser role in our experiment. Nonetheless, it is important to recognise that our estimates of ethnic discrimination may be subject to positive bias and, therefore conservative. Note, however, that social desirability bias is unlikely to affect recruiters' judgement with regard to factors that mitigate ethnic discrimination, which are the main focus of our study.

Third, in our experiment, we stated that all fictitious applicants possessed a sufficiently high level of German language proficiency. This allowed us to avoid implausible combinations of applicant characteristics and focus on the effect of other skills less investigated by the literature. However, the magnitude of ethnic discrimination can increase at lower levels of language proficiency. This would imply that our estimates of ethnic discrimination are lower-bound estimates.

Lastly, our study focuses on the first stage of the hiring process, specifically the screening of applicants for a job interview. This focus is aligned with established research, which indicates that most instances of discrimination occur during the initial screening of applications (Neumark & Rich, 2019). Although our study does not extend beyond this initial phase to investigate events during and after job interviews, it provides valuable insights into an essential part of the hiring process during which discrimination is known to be significant.

References

- Agerström, Jens, Fredrik Björklund, Rickard Carlsson, and Dan-Olof Rooth. "Warm and Competent Hassan = Cold and Incompetent Eric: A Harsh Equation of Real-Life Hiring Discrimination". *Basic and Applied Social Psychology* 34 (2012), 359–366. DOI: [10.1080/01973533.2012.693438](https://doi.org/10.1080/01973533.2012.693438).
- Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer. "Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success". *Journal of Labor Economics* 34 (2016), S361–S401. DOI: [10.1086/682938](https://doi.org/10.1086/682938).
- Andernach, Björn and Reinhard Schunck. "Investigating the feasibility of a factorial survey in a CATI". In: 2014.
- Arellano-Bover, Jaime. "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size". *Journal of Labor Economics* (forthcoming). DOI: [10.1086/723500](https://doi.org/10.1086/723500).
- Arrow, Kenneth J. "The Theory of Discrimination". In: *Discrimination in Labor Markets*. Princeton University Press, 1973, 3–33. URL: <http://www.jstor.org/stable/j.ctt13x10hs.5> (visited on 07/04/2023).
- Auspurg, K. and T. Hinz. *Factorial Survey Experiments*. Sage, Thousand Oaks, CA, 2014. ISBN: 9781452274188. DOI: [10.4135/9781483398075](https://doi.org/10.4135/9781483398075).
- Baert, Stijn. "Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments Since 2005". In: *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Ed. by S. Michael Gaddis. Springer International Publishing, 2018, 63–77. DOI: [10.1007/978-3-319-71153-9_3](https://doi.org/10.1007/978-3-319-71153-9_3).
- Baert, Stijn, Andrea Albanese, Sofie du Gardein, Jolien Ovaere, and Jarno Stappers. "Does work experience mitigate discrimination?" *Economics Letters* 155 (2017), 35–38. ISSN: 0165-1765. DOI: [10.1016/j.econlet.2017.03.011](https://doi.org/10.1016/j.econlet.2017.03.011).
- Baert, Stijn and Ann-Sophie De Pauw. "Is ethnic discrimination due to distaste or statistics?" *Economics Letters* 125 (2014), 270–273. DOI: [10.1016/j.econlet.2014.09.020](https://doi.org/10.1016/j.econlet.2014.09.020).
- Baert, Stijn, Louis Lippens, and Hannah Van Borm. *Selecting Names for Experiments on Ethnic Discrimination*. IZA Discussion Paper No. 15524. IZA, 2022. DOI: [10.2139/ssrn.4206225](https://doi.org/10.2139/ssrn.4206225).
- Bertrand, M. and E. Duflo. "Field Experiments on Discrimination". In: *Handbook of Field Experiments*. Ed. by Abhijit Vinayak Banerjee and Esther Duflo. 1. Handbook of Economic Field Experiments. North-Holland, 2017, 309–393. DOI: [10.1016/bs.hefe.2016.08.004](https://doi.org/10.1016/bs.hefe.2016.08.004).
- Bertrand, Marianne and Sendhil Mullainathan. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labour Market Discrimination". *American Economic Review* 94 (2004), 991–1013. DOI: [10.1257/0002828042002561](https://doi.org/10.1257/0002828042002561).
- BIBB. "Datenreport zum Berufsbildungsbericht 2021. Informationen und Analysen zur Entwicklung der beruflichen Bildung." In: Bonn: Bundesinstitut für Berufsbildung (Hrsg.), 2021. URL: https://www.bibb.de/datenreport/de/datenreport_2021.php (visited on 08/15/2023).
- Biesma, R.G., M. Pavlova, G.G. van Merode, and W. Groot. "Using conjoint analysis to estimate employers preferences for key competencies of master level Dutch graduates entering the public health field". *Economics of Education Review* 26 (2007), 375–386. ISSN: 0272-7757. DOI: [10.1016/j.econedurev.2006.01.004](https://doi.org/10.1016/j.econedurev.2006.01.004).

- Blommaert, L., M. Coenders, and F. van Tubergen. "Ethnic Discrimination in Recruitment and Decision Makers' Features: Evidence from Laboratory Experiment and Survey Data using a Student Sample." *Social Indicators Research* 116 (2014), 731—754. DOI: [10.1007/s11205-013-0329-4](https://doi.org/10.1007/s11205-013-0329-4).
- Borghans, Lex, Bas Ter Weel, and Bruce A. Weinberg. "People Skills and the Labor-Market Outcomes of Underrepresented Groups". *ILR Review* 67 (2014), 287–334. DOI: [10.1177/001979391406700202](https://doi.org/10.1177/001979391406700202).
- Carlsson, Magnus and Dan-Olof Rooth. "Evidence of ethnic discrimination in the Swedish labor market using experimental data". *Labour Economics* 14 (2007), 716–729. ISSN: 0927-5371. DOI: [10.1016/j.labeco.2007.05.001](https://doi.org/10.1016/j.labeco.2007.05.001).
- Dahl, Malte. "Alike but Different: How Cultural Distinctiveness Shapes Immigrant-Origin Minorities' Access to the Labour Market." *Int. Migration & Integration* 23 (2022), 2269—2287. DOI: [10.1007/s12134-021-00844-y](https://doi.org/10.1007/s12134-021-00844-y).
- Damelang Andreas, Sabine Ebenspreger and Felix Stumpf. "Foreign Credential Recognition and Immigrants' Chances of being Hired for Skilled Jobs – Evidence from a Survey Experiment among Employers". *Social Forces* 49 (2020), 1–24. DOI: [10.1093/sf/soz154](https://doi.org/10.1093/sf/soz154).
- Deming, David J. "The Growing Importance of Social Skills in the Labor Market*". *The Quarterly Journal of Economics* 132 (2017), 1593–1640. ISSN: 0033-5533. DOI: [10.1093/qje/qjx022](https://doi.org/10.1093/qje/qjx022).
- European Commission. "Proposal for a Decision of the European Parliament and of the Council on a European Year of Skills 2023" (2022). URL: <https://ec.europa.eu/social/BlobServlet?docId=26148&langId=en>.
- European Labour Authority. *EURES Report on Labour Shortages and Surpluses*. Publications Office of the European Union, 2022. DOI: [10.2883/67288](https://doi.org/10.2883/67288).
- Friedrich, Anett, Christian Gerhards, Sabine Mohr, Klaus Troltsch, and Kathrin Weiß. "BIBB Training Panel - An Establishment Panel on Training and Competence Development 2011 to 2020 long. GWA_1.0." Research Data Center at BIBB (ed., data access); Bonn: Federal Institute for Vocational Education and Training (2022). DOI: [10.7803/371.1120.1.2.10](https://doi.org/10.7803/371.1120.1.2.10).
- Ganzeboom, Harry. *Occupation coding: Do's and Dont's. With special reference to the International Standard Classification of Occupations ISCO-88. With an Extension on ISCO-08*. Working paper. Free University Amsterdam, 2010. URL: <http://www.harryganzeboom.nl/isco08/index.htm>.
- Ganzeboom, Harry B. G. and Donald J. Treiman. "Three Internationally Standardised Measures for Comparative Research on Occupational Status". In: *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*. Ed. by Jürgen H. P. Hoffmeyer-Zlotnik and Christof Wolf. Boston, MA: Springer US, 2003, 159–193. ISBN: 978-1-4419-9186-7. DOI: [10.1007/978-1-4419-9186-7_9](https://doi.org/10.1007/978-1-4419-9186-7_9).
- Guryan, Jonathan and Kerwin Kofi Charles. "Taste-Based or Statistical Discrimination: the Economics of Discrimination Returns to its Roots". *The Economic Journal* 123 (2013), F417–F432. ISSN: 00130133, 14680297. URL: <http://www.jstor.org/stable/42919257> (visited on 08/09/2023).
- Hagendoorn, Louk. "Intergroup Biases in Multiple Group Systems: The Perception of Ethnic Hierarchies". *European Review of Social Psychology* 6 (1995), 199–228. DOI: [10.1080/1479277943000058](https://doi.org/10.1080/1479277943000058).

- Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. "Validating vignette and conjoint survey experiments against real-world behavior". *Proceedings of the National Academy of Sciences* 112 (2015), 2395–2400. DOI: [10.1073/pnas.1416587112](https://doi.org/10.1073/pnas.1416587112).
- Heckman, James J., Jora Stixrud, and Sergio Urzua. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior". *Journal of Labor Economics* 24 (2006), 411–482. DOI: [10.1086/504455](https://doi.org/10.1086/504455).
- Heublein, U., Hutzsch C., Isleib S., König R., Richter J., and Woisch A. "Ursachen des Studienabbruchs und beruflicher Verbleib von Studienabbrechern. [Causes of university dropout and job retention]". In: Hannover, Germany: DZHW, 2017.
- Heublein, Ulrich. "Student Drop-out from German Higher Education Institutions". *European Journal of Education* 49 (2014), 497–513. DOI: [10.1111/ejed.12097](https://doi.org/10.1111/ejed.12097).
- Hamburg, Martin and Rolf van der Velden. "Skills and Graduate Recruitment Process: Evidence from Two Discrete Choice Experiments". *Economics of Education Review* 49 (2015), 24–41. DOI: [10.1016/j.econedurev.2015.07.001](https://doi.org/10.1016/j.econedurev.2015.07.001).
- Kaas, Leo and Christian Manger. "Ethnic Discrimination in Germany's Labour Market: A Field Experiment". *German Economic Review* 13 (2012), 1–20. DOI: [10.1111/j.1468-0475.2011.00538.x](https://doi.org/10.1111/j.1468-0475.2011.00538.x).
- Kline, Patrick, Evan K Rose, and Christopher R Walters. "Systemic Discrimination Among Large U.S. Employers". *The Quarterly Journal of Economics* 137 (2022), 1963–2036. ISSN: 0033-5533. DOI: [10.1093/qje/qjac024](https://doi.org/10.1093/qje/qjac024).
- Kline, Patrick and Christopher Walters. "Reasonable Doubt: Experimental Detection of Job-Level Employment Discrimination". *Econometrica* 89 (2021), 765–792. DOI: [10.3982/ECTA17489](https://doi.org/10.3982/ECTA17489).
- Kubler, Dorothea, Julia Schmid, and Robert Stuber. "Gender Discrimination in Hiring across Occupations: a Nationally-representative Vignette Study". *Labour Economics* 55 (2018), 215–229. DOI: [10.1016/j.labeco.2018.10.002](https://doi.org/10.1016/j.labeco.2018.10.002).
- Lahey, Joanna N. and Douglas R. Oxley. "Discrimination at the Intersection of Age, Race, and Gender: Evidence from an Eye-Tracking Experiment". *Journal of Policy Analysis and Management* 40 (2021), 1083–1119. DOI: [10.1002/pam.22281](https://doi.org/10.1002/pam.22281).
- Leys, Christophe, Christophe Ley, Olivier Klein, Philippe Bernard, and Laurent Licata. "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median". *Journal of Experimental Social Psychology* 49 (2013), 764–766. ISSN: 0022-1031. DOI: [10.1016/j.jesp.2013.03.013](https://doi.org/10.1016/j.jesp.2013.03.013).
- Lippens, Louis, Siel Vermeiren, and Stijn Baert. "The state of hiring discrimination: A meta-analysis of (almost) all recent correspondence experiments". *European Economic Review* 151 (2023), 104315. ISSN: 0014-2921. DOI: [10.1016/j.euroecorev.2022.104315](https://doi.org/10.1016/j.euroecorev.2022.104315).
- McGrath, John. *Report on Labour Shortages and Surpluses*. European Labor Authority, 2021. URL: <https://www.ela.europa.eu/en/media/725>.
- Mergener, Alexandra and Tobias Maier. "Immigrants' Chances of Being Hired at Times of Skill Shortages: Results from a Factorial Survey Experiment Among German Employers". *Journal of International Migration and Integration* 20 (2019), 155–177. DOI: [10.1007/s12134-018-0598-5](https://doi.org/10.1007/s12134-018-0598-5).
- Neumark, David. "Experimental Research on Labor Market Discrimination". *Journal of Economic Literature* 56 (2018), 799–866. DOI: [10.1257/jel.20161309](https://doi.org/10.1257/jel.20161309).
- Neumark, David and Judith Rich. "Do Field Experiments on Labor and Housing Markets Overstate Discrimination? A Re-examination of the Evidence". *ILR Review* 72 (2019), 223–252. DOI: [10.1177/0019793918759665](https://doi.org/10.1177/0019793918759665).

- Nunley, John M., Adam Pugh, Nicholas Romero, and R. Alan Seals. "College major, internship experience, and employment opportunities: Estimates from a résumé audit". *Labour Economics* 38 (2016), 37–46. ISSN: 0927-5371. DOI: [10.1016/j.labeco.2015.11.002](https://doi.org/10.1016/j.labeco.2015.11.002).
- Oreopoulos, Philip. "Why Do Skilled Immigrants Struggle in the Labour Market? A Field Experiment with Thirteen Thousand Resumes". *American Economic Journal: Economic Policy* 3 (2011), 148–171. DOI: [10.1257/pol.3.4.148](https://doi.org/10.1257/pol.3.4.148).
- Phelps, Edmund S. "The Statistical Theory of Racism and Sexism". *The American Economic Review* 62 (1972), 659–661. ISSN: 00028282. URL: <http://www.jstor.org/stable/1806107> (visited on 07/04/2023).
- Phillips, David C. "Do Comparisons of Fictional Applicants Measure Discrimination When Search Externalities are Present? Evidence from Existing Experiments". *The Economic Journal* 129 (2019), 2240–2264. ISSN: 0013-0133. DOI: [10.1111/ecoj.12628](https://doi.org/10.1111/ecoj.12628).
- Piopiunik, Marc, Guido Schwerdt, Lisa Simon, and Ludger Woessmann. "Skills, signals, and employability: An experimental investigation". *European Economic Review* 123 (2020), 103374. ISSN: 0014-2921. DOI: [10.1016/j.eurocorev.2020.103374](https://doi.org/10.1016/j.eurocorev.2020.103374).
- Protsch, Paula and Heike Solga. "How Employers Use Signals of Cognitive and Noncognitive Skills at Labour Market Entry: Insights from Field Experiments". *European Sociological Review* 31 (2015), 521–532. DOI: [10.1093/esr/jcv056](https://doi.org/10.1093/esr/jcv056).
- Protsch, Paula and Heike Solga. "Going across Europe for an Apprenticeship? A Factorial Survey Experiment on Employers' Hiring Preferences in Germany". *Journal of European Social Policy* 27 (2017), 387–399. DOI: [10.1177/0958928717719200](https://doi.org/10.1177/0958928717719200).
- Quillian, Lincoln, Pager Devah, Hexel Ole, and Midtbøen Arnfinn. "Meta-Analysis of Field Experiments Shows No Change in Racial Discrimination in Hiring over Time". *Proceedings of the National Academy of Sciences* 114 (2017), 10870–10875. DOI: [10.1073/pnas.1706255114](https://doi.org/10.1073/pnas.1706255114).
- Quillian, Lincoln, John J Lee, and Mariana Oliver. "Evidence from Field Experiments in Hiring Shows Substantial Additional Racial Discrimination after the Callback". *Social Forces* 99 (2020), 732–759. ISSN: 0037-7732. DOI: [10.1093/sf/soaa026](https://doi.org/10.1093/sf/soaa026).
- Sidanis, J. and F Pratto. *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge: Cambridge University Press, 1999. DOI: [10.1017/CBO9781139175043](https://doi.org/10.1017/CBO9781139175043).
- Thijssen, L., M. Coenders, and B. Lancee. "Ethnic Discrimination in the Dutch Labor Market: Differences Between Ethnic Minority Groups and the Role of Personal Information About Job Applicants—Evidence from a Field Experiment." *Journal of International Migration and Integration* 22 (2021), 1125–1150. DOI: [10.1007/s12134-020-00795-w](https://doi.org/10.1007/s12134-020-00795-w).
- Tobback, Ilse, Dieter Verhaest, and Stijn Baert. "Student Access to Apprenticeships: Evidence from a Vignette Experiment". *Industrial Relations: A Journal of Economy and Society* 59 (2020), 435–465. DOI: [10.1111/irel.12258](https://doi.org/10.1111/irel.12258).
- Van Borm, Hannah, Louis Lippens, and Stijn Baert. *An Arab, an Asian, and a Black Guy Walk into a Job Interview: Ethnic Stigma in Hiring after Controlling for Social Class*. IZA Discussion Paper No. 15707. IZA, 2022. DOI: [10.2139/ssrn.4281262](https://doi.org/10.2139/ssrn.4281262).

Appendix: Tables and Figures

Table A1

Descriptive statistics.

	Total sample (1)	German surname (2)	Non-German surname (3)	Difference (2)-(3) (4)
A. Recruiter characteristics				
Gender: female	0.449	0.442	0.452	-0.009
Age: below mean (<47)	0.439	0.435	0.440	-0.005
Risk attitude (0-10 scale)	5.685	5.698	5.681	0.017
B. Firm characteristics				
Location: East Germany	0.198	0.198	0.198	0.000
Share female employees	0.398	0.397	0.399	-0.002
Share employees with migration background	0.149	0.148	0.149	-0.001
<i>Firm size (number of employees):</i>				
1-19	0.168	0.167	0.168	-0.002
20-99	0.270	0.267	0.271	-0.004
100-199	0.221	0.219	0.221	-0.003
200-499	0.238	0.238	0.238	-0.000
above 500	0.104	0.110	0.102	0.008
<i>Industry:</i>				
Agriculture, Forestry & Mining	0.049	0.048	0.049	-0.000
Manufacturing	0.323	0.320	0.324	-0.004
Construction	0.063	0.064	0.063	0.001
Trade & Repair	0.150	0.155	0.149	0.006
Business services	0.156	0.157	0.155	0.002
Other services	0.125	0.120	0.127	-0.006
Medical & Nursing services	0.042	0.041	0.041	-0.000
Public service & Education	0.093	0.095	0.093	0.002
C. Evaluation outcome				
Interview probability	6.76	6.87	6.72	-0.150***
Observations	11,748	2,918	8,830	

Notes: The table shows the means of variables for the total vignette sample and for sub-samples of fictitious candidates with German and foreign surnames separately. All variables except *Interview probability score*, which takes values 1-10, and *Risk attitude*, which takes values from 0 (low risk) to 10 (high risk), are binary. Column (4) presents the difference in means between German and Foreign samples and shows a significance level from the t-test of mean difference. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

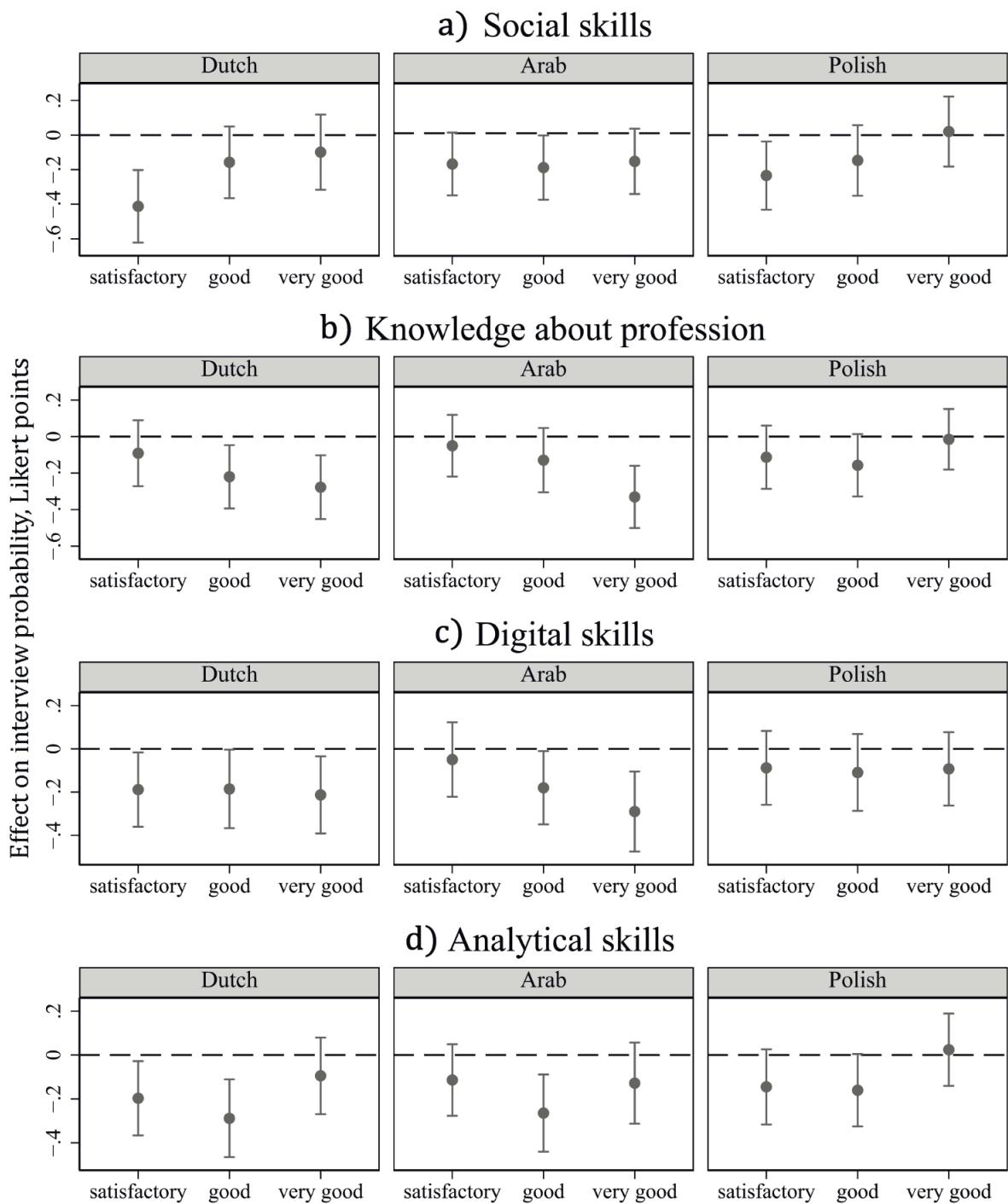


Fig. A1. Interaction between skills and ethnicity, marginal effects.

Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure plots evaluations of applicants with ethnicity-specific non-German and German surnames at different levels of skills and professional knowledge. Estimates are obtained by calculation of margins from a RE regression of the applicant evaluation on ethnicity, the corresponding skill attribute, their interactions, all other applicant characteristics, question order, and firm- and recruiter-level controls. Standard errors are clustered at the recruiter level. 95% confidence intervals are shown on the graphs.

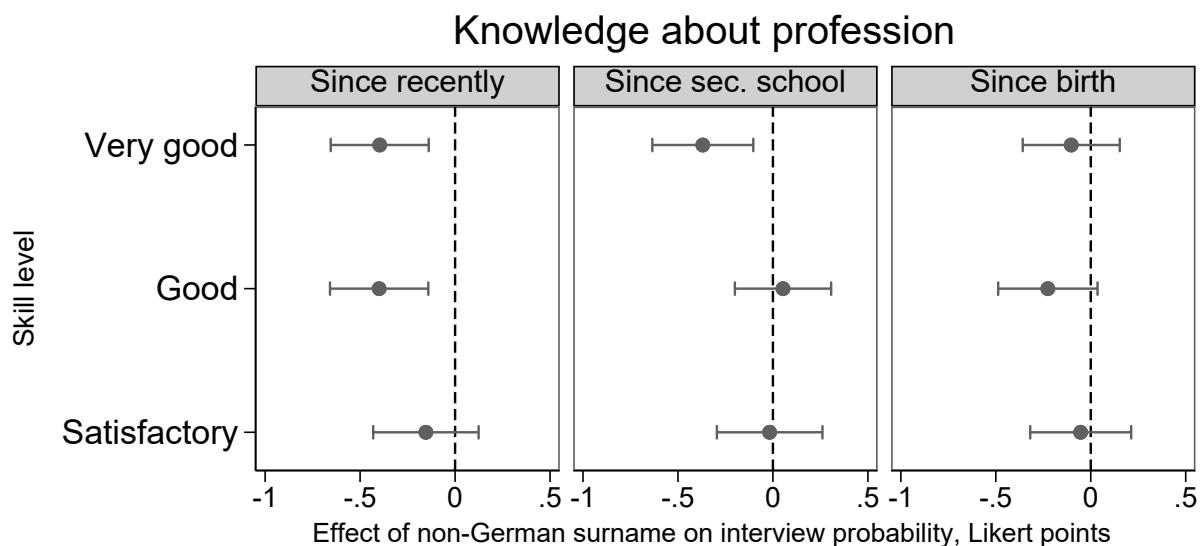


Fig. A2. Interaction effects between non-German ethnicity and knowledge about profession depending on the time of residence.

Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure plots the marginal effects of having a non-German surname at different levels of knowledge about the profession depending on the time of residence. Estimates are obtained by calculation of marginal effects from a regression of the applicant evaluation on a three-way interaction between non-German ethnicity, knowledge about profession and time of residence, all applicant's characteristics, question order, firm- and recruiter-level controls. Standard errors are clustered at the recruiter level. 95% confidence intervals are shown on the graphs.

Table A2

Interactions with occupation, firm and recruiter characteristics.

	Interview probability				
	(1)	(2)	(3)	(4)	(5)
Firm: size < 50 employees	-0.231*** (0.071)	-0.107 (0.097)	-0.231*** (0.071)	-0.231*** (0.071)	-0.096 (0.098)
Firm: in East Germany	-0.157* (0.084)	-0.117 (0.113)	-0.156* (0.084)	-0.157* (0.084)	-0.110 (0.114)
Firm: share migr. employees < mean	-0.112 (0.070)	-0.130 (0.095)	-0.112 (0.070)	-0.112 (0.070)	-0.131 (0.096)
Firm: size < 50 employees × non-German		-0.164* (0.092)			-0.178* (0.095)
Firm: in East Germany × non-German			-0.052 (0.105)		-0.060 (0.106)
Firm: share migr. employees < mean × non-German		0.023 (0.089)			0.024 (0.090)
Recruiter: female	0.049 (0.064)	0.049 (0.064)	0.097 (0.088)	0.048 (0.064)	0.106 (0.089)
Recruiter: age < 45	0.159** (0.063)	0.159** (0.063)	0.073 (0.087)	0.159** (0.063)	0.075 (0.088)
Recruiter: risk attitude < mean	0.189*** (0.064)	0.189*** (0.064)	0.110 (0.088)	0.189*** (0.064)	0.110 (0.088)
Recruiter: female × non-German			-0.064 (0.084)		-0.077 (0.085)
Recruiter: age < 45 × non-German			0.114 (0.083)		0.111 (0.084)
Recruiter: risk attitude < mean × non-German			0.106 (0.084)		0.105 (0.084)
Occupation: Communicate with clients	-0.059 (0.067)	-0.058 (0.067)	-0.059 (0.067)	-0.053 (0.095)	-0.034 (0.095)
Occupation: SIOPS q1	0.288*** (0.097)	0.288*** (0.097)	0.289*** (0.097)	0.347** (0.139)	0.328** (0.139)
Occupation: SIOPS q2	0.247*** (0.091)	0.247*** (0.091)	0.248*** (0.091)	0.452*** (0.128)	0.450*** (0.128)
Occupation: SIOPS q3	0.040 (0.092)	0.039 (0.092)	0.039 (0.092)	0.209 (0.135)	0.225* (0.135)
Occupation: Communicate with clients × non-German				-0.008 (0.089)	-0.033 (0.088)
Occupation: SIOPS q1 × non-German				-0.079 (0.131)	-0.053 (0.132)
Occupation: SIOPS q2 × non-German				-0.272** (0.124)	-0.269** (0.124)
Occupation: SIOPS q3 × non-German				-0.226* (0.128)	-0.248* (0.129)
(Occupation: SIOPS q4 - reference level)					
(All applicant's characteristics included as controls)					
Observations	11,748	11,748	11,748	11,748	11,748

Notes: The coefficients and standard errors are obtained from RE regressions of the evaluation of job interview probability on all applicant's characteristics, order controls and listed occupation, firm and recruiter characteristics and interaction terms between these occupation, firm and recruiter characteristics and non-German ethnicity. Standard errors clustered at the respondent level (1958 clusters) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Supplementary Materials

for "*Can skills alleviate ethnic discrimination in hiring? Evidence from the German apprenticeship market.*"
(for online publication)

Contents

A Additional Tables and Figures	S-2
B Experiment text, English translation	S-10
B.1 Handout materials	S-10
B.2 Experiment introduction	S-11
B.3 Example task	S-12
C Experiment text, German versions	S-13
C.1 Handout materials (German)	S-13
C.2 Experiment introduction (German)	S-14
C.3 Example task (German)	S-15

List of Tables

S1 Overview of empirical studies on factors mitigating ethnic discrimination in hiring.	S-2
S2 The effect of applicant characteristics on access to apprenticeships. Alternative regression specifications.	S-3
S3 The effect of applicant characteristics on access to apprenticeships.	S-4
S4 Interaction effects between skills and ethnicity (Table 2 extended).	S-5

List of Figures

S1 Distribution of the outcome variable (interview probability).	S-6
S2 Evaluations of German and non-German applicants by skill level.	S-7
S3 Ethnic differences in evaluations by occupational status measured by SIOPS and ISEI indexes.	S-8
S4 Interactions between skills, non-German ethnicity and firm, recruiter and occupation characteristics, marginal effects.	S-9

A Additional Tables and Figures

Table S1

Overview of empirical studies on factors mitigating ethnic discrimination in hiring.

Publication	Country	Ethnicities included	Factors studied	Effect
Oreopoulos (2011)	Canada	Chinese, Indian, Pakistani, British, Greek	Language fluency Local work experience Multinational firm experience Education from selective schools Extracurricular activities	(+) (+) (0) (0) (0)
Agerström et al. (2012)	Sweden	Arab	Self-reported information on personality traits (Warmth & Competence)	(+)
Kaas and Manger (2012)	Germany	Turkish	Reference letters	(+)
Baert et al. (2017)	Belgium	Turkish, Moroccan, Slovakian, Ghanaian	Work experience	(+)
Thijssen et al. (2021a)	Netherlands	Western European, South Asian, Eastern European, Asian, Latin American, Central African, and MENA origin	Self-reported favourable information about hard and soft skills	(0)
Thijssen et al. (2021b)	Netherlands Germany	Turkish	Picture in the CV Past work performance GPA	(0) (0) (0)
Dahl (2022)	Denmark	Arab	Liberal lifestyle Civic participation	(0) (0)

Notes: The labels (+) and (0) are used to indicate factors that were discovered to reduce ethnic disparities in hiring and factors that had no substantial impact on ethnic disparities in hiring, respectively.

Table S2

The effect of applicant characteristics on access to apprenticeships. Alternative regression specifications.

	Interview probability				
	(1)	(2)	(3)	(4)	(5)
<i>Ethnicity (reference level - German)</i>					
Dutch	-0.196*** (0.051)	-0.196*** (0.051)	-0.193*** (0.052)	-0.193*** (0.052)	-0.197*** (0.051)
Arab	-0.172*** (0.051)	-0.172*** (0.051)	-0.172*** (0.052)	-0.172*** (0.052)	-0.172*** (0.052)
Polish	-0.097** (0.049)	-0.096** (0.049)	-0.092* (0.050)	-0.087* (0.050)	-0.100** (0.049)
<i>Skills (reference level - Satisfactory)</i>					
Social skills, very good	1.138*** (0.045)	1.138*** (0.045)	1.136*** (0.045)	1.137*** (0.045)	1.138*** (0.045)
Social skills, good	0.993*** (0.045)	0.993*** (0.045)	0.989*** (0.046)	0.989*** (0.046)	0.995*** (0.046)
Knowl. about profession, very good	1.112*** (0.046)	1.113*** (0.046)	1.111*** (0.047)	1.114*** (0.047)	1.112*** (0.046)
Knowl. about profession, good	0.933*** (0.045)	0.932*** (0.045)	0.932*** (0.046)	0.930*** (0.046)	0.933*** (0.045)
Digital skills, very good	0.464*** (0.041)	0.464*** (0.041)	0.465*** (0.041)	0.465*** (0.041)	0.464*** (0.041)
Digital skills, good	0.443*** (0.042)	0.443*** (0.042)	0.449*** (0.042)	0.449*** (0.042)	0.441*** (0.042)
Analytical skills, very good	0.486*** (0.041)	0.486*** (0.041)	0.490*** (0.042)	0.489*** (0.042)	0.485*** (0.041)
Analytical skills, good	0.464*** (0.043)	0.464*** (0.043)	0.464*** (0.043)	0.464*** (0.043)	0.464*** (0.043)
Female	-0.076* (0.040)	-0.076* (0.040)	-0.080** (0.040)	-0.080** (0.040)	-0.074* (0.040)
<i>Time in Germany (reference level - Since birth)</i>					
Since recently	-0.372*** (0.045)	-0.372*** (0.045)	-0.364*** (0.045)	-0.363*** (0.045)	-0.376*** (0.045)
Since secondary school	-0.064 (0.041)	-0.064 (0.041)	-0.065 (0.042)	-0.065 (0.042)	-0.064 (0.042)
<i>Education (reference level - Lower secondary)</i>					
General secondary	1.343*** (0.057)	1.342*** (0.057)	1.352*** (0.058)	1.350*** (0.058)	1.338*** (0.057)
Higher secondary	1.327*** (0.065)	1.327*** (0.065)	1.333*** (0.067)	1.333*** (0.067)	1.325*** (0.065)
Unfinished higher education	0.819*** (0.063)	0.818*** (0.063)	0.830*** (0.065)	0.827*** (0.065)	0.814*** (0.063)
Model	RE	RE	OLS	OLS	FE
Respondent & firm controls	✓	-	✓	-	-
Observations	11,748	11,748	11,748	11,748	11,748

Notes: Estimates in Columns (1)-(2), (3)-(4) and (5) were obtained from the random effects, OLS and fixed effects regressions of interview probability on the applicant characteristics and question order dummies, respectively. Regressions in Columns (1) and (3) additionally include respondent's age, gender and risk attitude, and firm's size, industry and location as controls. Standard errors clustered at the respondent level (1958 clusters) are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table S3

The effect of applicant characteristics on access to apprenticeships.

	Interview probability	
	(1)	(2)
<i>Ethnicity (reference level - German)</i>		
non-German	-0.155*** (0.041)	
Dutch		-0.196*** (0.051)
Arab		-0.172*** (0.051)
Polish		-0.097** (0.049)
<i>Skills (reference level - Satisfactory)</i>		
Social skills, very good	1.138*** (0.045)	1.138*** (0.045)
Social skills, good	0.994*** (0.045)	0.993*** (0.045)
Knowl. about profession, very good	1.112*** (0.046)	1.112*** (0.046)
Knowl. about profession, good	0.934*** (0.045)	0.933*** (0.045)
Digital skills, very good	0.464*** (0.041)	0.464*** (0.041)
Digital skills, good	0.444*** (0.042)	0.443*** (0.042)
Analytical skills, very good	0.486*** (0.041)	0.486*** (0.041)
Analytical skills, good	0.464*** (0.043)	0.464*** (0.043)
Female	-0.076* (0.040)	-0.076* (0.040)
<i>Time in Germany (reference level - Since birth)</i>		
Since recently	-0.373*** (0.045)	-0.372*** (0.045)
Since secondary school	-0.065 (0.041)	-0.064 (0.041)
<i>Education (reference level - Lower secondary)</i>		
General secondary	1.342*** (0.057)	1.343*** (0.057)
Higher secondary	1.326*** (0.065)	1.327*** (0.065)
Unfinished higher education	0.818*** (0.063)	0.819*** (0.063)
Observations	11,748	11,748

Notes: The coefficients in Columns (1)-(2) and the standard errors are obtained from RE regressions described by Equation 1. All models include question order dummies, respondent's age, gender and risk attitude, and firm's size, industry and location as controls. Standard errors clustered at the respondent level (1958 clusters) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

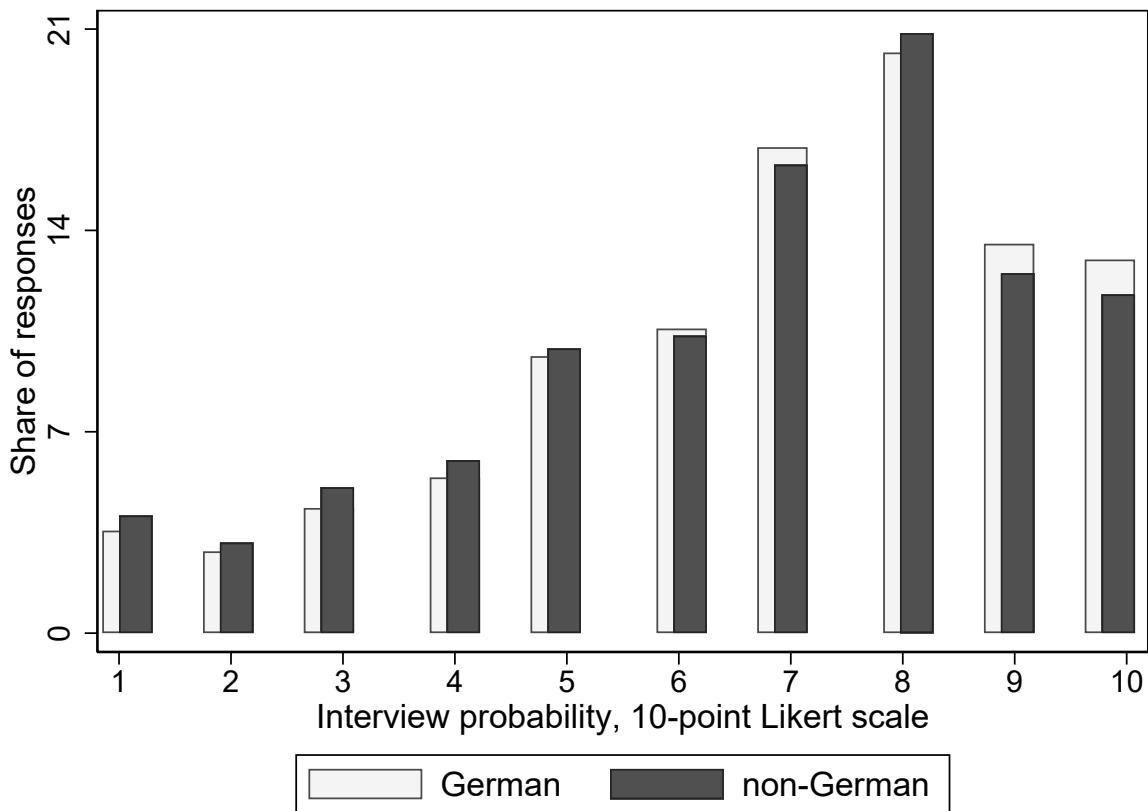
Table S4

Interaction effects between skills and ethnicity (Table 2 extended).

	(1)	(2)
Social skills × non-German	0.103** (0.052)	
Very good × non-German		0.204** (0.100)
Good × non-German		0.151 (0.107)
Knowl. about profession × non-German	-0.108** (0.054)	
Very good × non-German		-0.215** (0.107)
Good × non-German		-0.113 (0.111)
Digital skills × non-German	-0.045 (0.052)	
Very good × non-German		-0.089 (0.104)
Good × non-German		-0.047 (0.102)
Analytical skills × non-German	0.041 (0.050)	
Very good × non-German		0.085 (0.101)
Good × non-German		-0.082 (0.098)
Non-German	-0.142 (0.211)	-0.109 (0.126)
Social skills, Very good	1.174*** (0.053)	1.039*** (0.089)
Social skills, Good	1.011*** (0.047)	0.933*** (0.087)
Knowl. about profession, Very good	1.080*** (0.052)	1.203*** (0.088)
Knowl. about profession, Good	0.919***	1.003***
Digital skills, Very good	0.441*** (0.049)	0.523*** (0.088)
Digital skills, Good	0.433*** (0.044)	0.478*** (0.088)
Analytical skills, Very good	0.506*** (0.049)	0.421*** (0.086)
Analytical skills, Good	0.474*** (0.045)	0.524*** (0.085)
(base level - Satisfactory)	(0.047)	(0.091)
Time in Germany, Since recently	-0.377*** (0.045)	-0.377*** (0.045)
Time in Germany, Since secondary school (base level - Since birth)	-0.067 (0.041)	-0.066 (0.041)
Education, General secondary	1.341*** (0.057)	1.341*** (0.057)
Education, Higher secondary	1.325*** (0.065)	1.325*** (0.065)
Education, Unfinished higher education (base level - Lower secondary)	0.818*** (0.063)	0.819*** (0.063)
Female	-0.077* (0.040)	-0.076* (0.040)
Observations	11,748	11,748

Notes: The table presents coefficients β_1' for interaction terms between skills and non-German ethnicity from Equation 2. Regressions include firm size, industry, location, recruiter's age and gender, and question order dummies as controls. The model in Column (1) shows interaction effects between skill level and non-German ethnicity assuming a linear effect of skill quality; the model in Column (2) presents interaction effects at each level of skill separately. Standard errors clustered at the recruiter level (1958 clusters) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

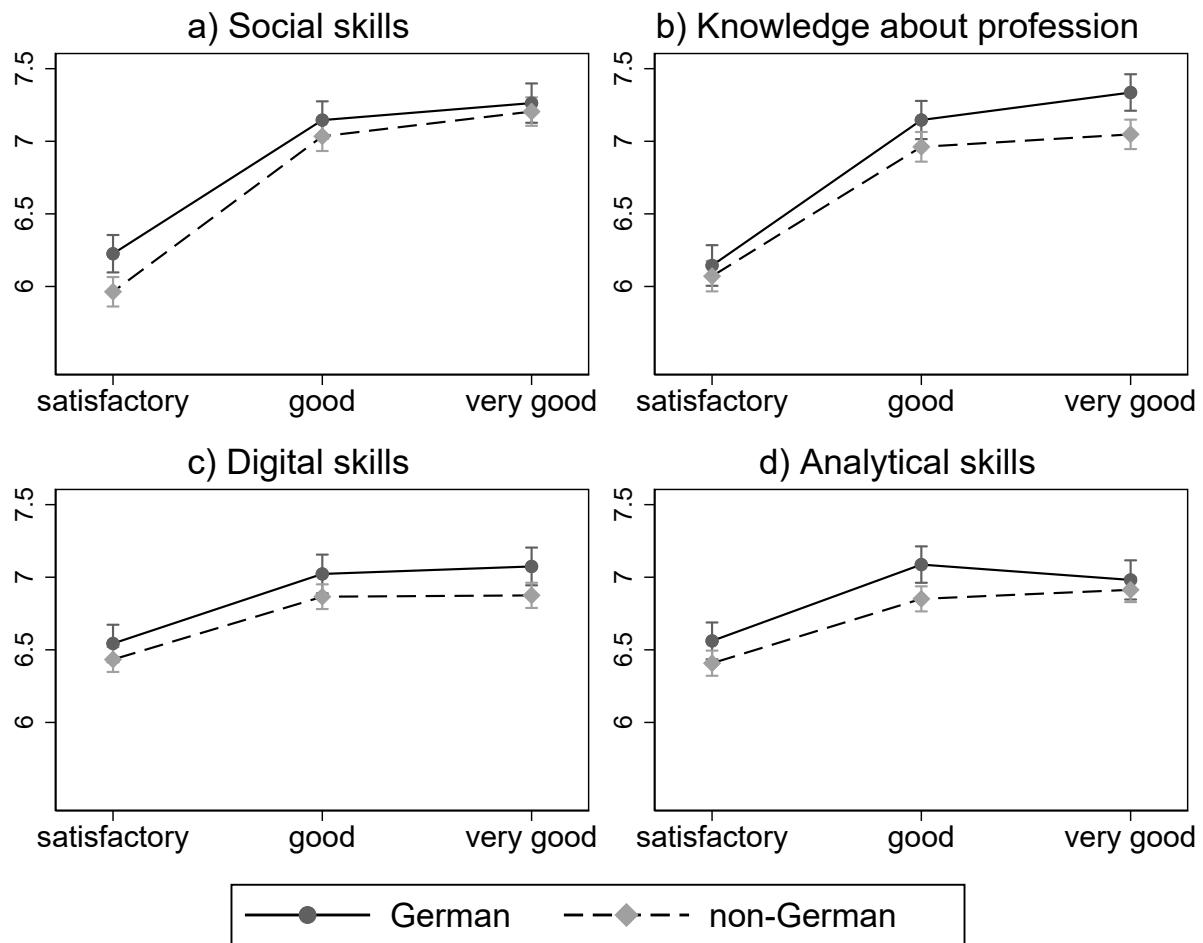
Fig. S1. Distribution of the outcome variable (interview probability).



Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The plot shows the fraction of each evaluation of the probability of invitation to a job interview for German and non-German applicants. Evaluations could be made on a scale from 1 to 10. The number of observations is 11,748 and was received from 1,958 recruiters.

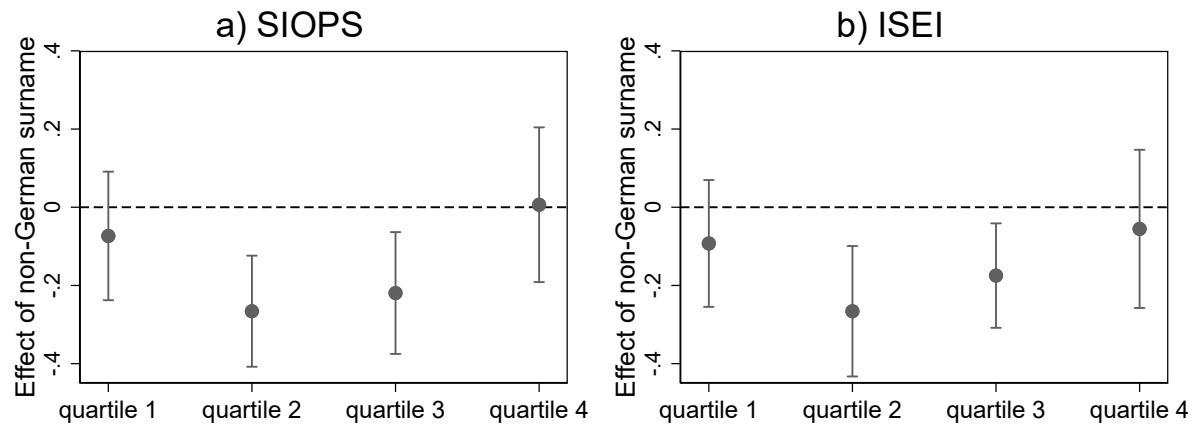
Fig. S2. Evaluations of German and non-German applicants by skill level.



Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure plots predicted evaluations of the job interview probability for German and non-German applicants depending on a skill level. Evaluation scores are measured on a 10-point Likert scale and are estimated from the regression model described in Equation 2. 95% confidence intervals are shown on the graphs.

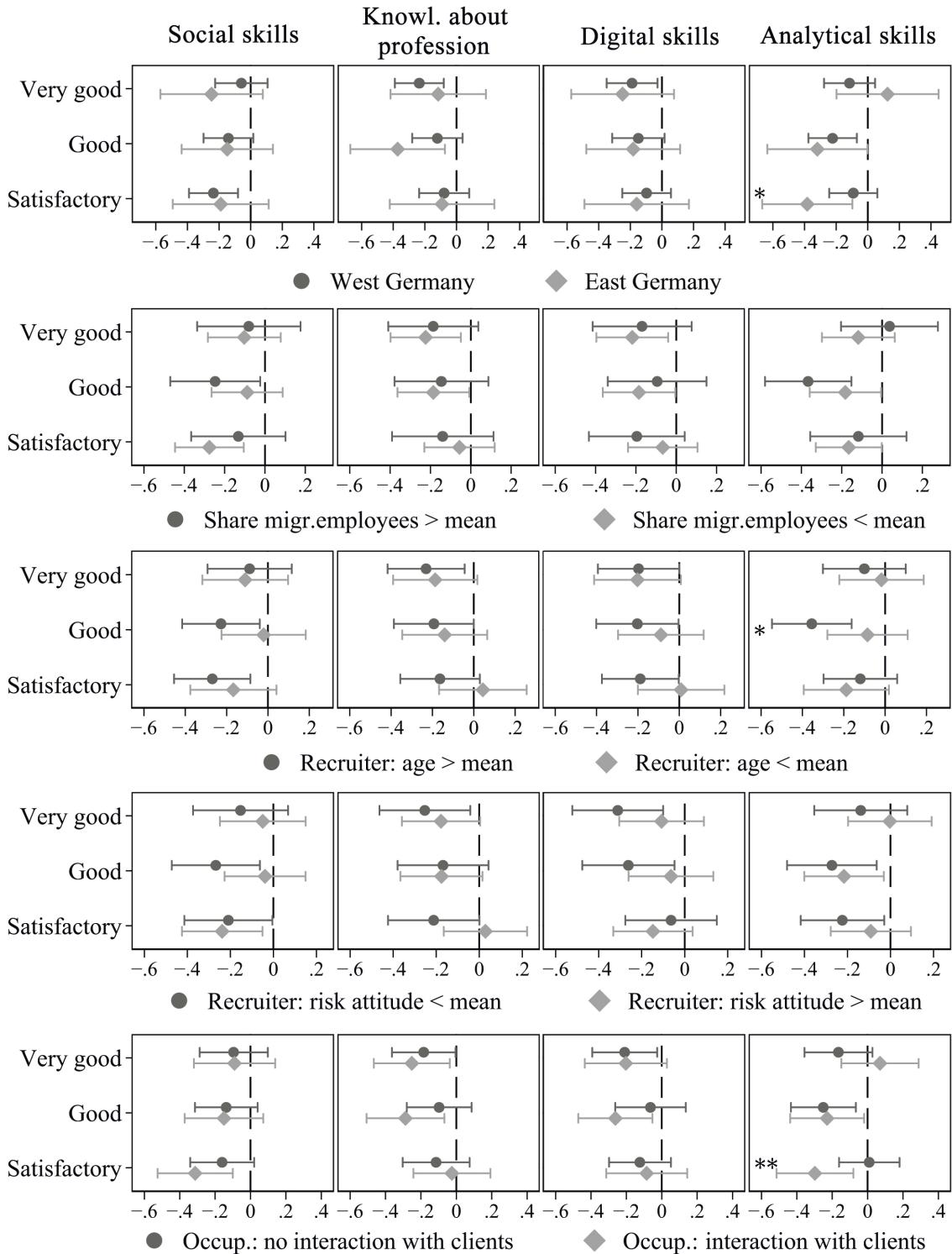
Fig. S3. Ethnic differences in evaluations by occupational status measured by SIOPS and ISEI indexes.



Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure presents the effect of having a non-German surname obtained from a regression of the applicant evaluation on ethnicity, occupation status index (SIOPS in a) and ISEI in b)), their interactions, all other applicant characteristics listed in Equation 1, as well as question order, firm- and recruiter-level controls. Standard errors are clustered at the recruiter level. 95% confidence intervals are shown on the graphs.

Fig. S4. Interactions between skills, non-German ethnicity and firm, recruiter and occupation characteristics, marginal effects.



Source: BIBB Qualification Panel 2021, authors' calculations. N = 11,748.

Notes: The figure plots marginal effects for non-German applicants at different levels of skills and knowledge about the profession relative to comparable applicants with German surnames. The effects on the x-axes are measured on a 10-point Likert scale. Estimates are obtained by calculating marginal effects from a regression with three-way interactions between skills, ethnicity, and firm, recruiter, and occupational characteristics. Standard errors are clustered at the respondent level and 95% confidence intervals are shown on the graphs. The stars on the left side of the plots show the significance level for the difference in estimates by the corresponding firm, recruiter and occupational characteristic: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B Experiment text, English translation

B.1 Handout materials

Please imagine that you are supposed to hire an applicant for the apprenticeship position in the occupation that is most frequently trained for in your company. Below we will present you with brief descriptions of six (6) fictitious applicants who applied for the position.

Each of these people is described by the following characteristics:

Surname

Gender

Time of residence in Germany

the highest **Education level** achieved

Please also imagine that, based on the application documents, you assess the following knowledge and skills of the applicants with grades (1 (very good); 2 (good), 3 (satisfactory)):

Social skills - an ability to work in teams, communicate and cooperate effectively, and cultural awareness, signalled by the extracurricular activities listed in the CV, such as team sports and volunteering experience, or via school grades for communication skills and teamwork.

Knowledge about profession - awareness of job requirements and job-related experience, which can be signalled by internships, professional orientation courses, side jobs and relevant hobbies.

Digital skills - the use of relevant software, computer and programming skills, signalled, for instance, by IT competencies reported in the CV and grades for the Informatics school course.

Analytical skills - reflectiveness, an ability to search for and process information, as well as an ability to approach problems from different angles, signalled by academic performance, hobbies and extracurricular activities, such as, for example, playing chess.

Please also imagine that the applicants only differ in the attributes mentioned, but not in all other characteristics. This means, for example, that all applicants have equally good knowledge of the German language and live close to your company. Applicants' school qualifications may also have been acquired abroad. They then correspond to the German school qualifications.

Based on these characteristics, we will ask you to indicate how likely this each applicant will get an invitation to a job interview in your company.

very unlikely

1 2 3 4 5 6 7 8 9 10

very likely

don't know

B.2 Experiment introduction

Regardless of the current situation in your company, please imagine that your company is urgently seeking apprentices, and you have six applications. I will introduce you to the applicants individually and ask you to rate how likely you would be to invite these people to an interview.

Please assume that the applicants only differ in the characteristics, knowledge and skills that I will read to you in each case. There are no differences in all other characteristics.

Please look at the sheet with the Handout materials.

(*for INTERVIEWER: Make sure that respondents have the sheet.*)

All the characteristics, knowledge and skills in which the applicants differ are listed there. That would be...

Surname

Gender

Time of residence in Germany

the highest **Education level** achieved

Social skills - an ability to work in teams, communicate and cooperate effectively, and cultural awareness.

Knowledge about profession - awareness of job requirements and job-related experience.

Digital skills - the use of relevant software, computer and programming skills.

Analytical skills - reflectiveness, an ability to search for and process information and to approach problems from different angles.

Based on the application documents, you have assessed the aforementioned knowledge and skills of the applicants with grades (1 (very good); 2 (good), 3 (satisfactory))

B.3 Example task

Please imagine that your firm has an urgent need to fill in an apprenticeship position in occupation [Insert: relevant for the firm occupation (based on pre-survey)].

I will describe six applicants for this position. The applicants differ only in attributes mentioned in the descriptions and are similar across any other characteristics. For each applicant please indicate how likely this person will get an invitation to a job interview.

Applicant 1:

The person's surname is Weber, it is a man, who lives in Germany since the start of secondary school, has finished General secondary education, and has very good knowledge about profession, very good digital skills, very good analytical skills, and satisfactory social skills.

On a scale from 1 to 10 what is the probability that this person will get an invitation to a job interview in your firm?

1 stands for „very unlikely“, 10 stands for „very likely“.

very unlikely										very likely
<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	<input type="checkbox"/> 6	<input type="checkbox"/> 7	<input type="checkbox"/> 8	<input type="checkbox"/> 9	<input type="checkbox"/> 10	<input type="checkbox"/> <i>don't know</i>

C Experiment text, German versions

C.1 Handout materials (German)

Bitte stellen Sie sich nun vor, dass Sie eine/n Bewerber/in für den Ausbildungsberuf einstellen sollen, der in Ihrem Betrieb am häufigsten ausgebildet wird. Im Folgenden werden wir Ihnen Kurzbeschreibungen von sechs (6) fiktiven Bewerbern/Bewerberinnen präsentieren, die sich auf die Stelle beworben haben.

Jede dieser Personen wird durch folgende Merkmale charakterisiert:

Nachname

Geschlecht

Aufenthaltsdauer in Deutschland

den höchsten **Bildungsabschluss**

Bitte stellen Sie sich ebenfalls vor, dass Sie auf Basis der Bewerbungsunterlagen die folgenden Kenntnisse und Kompetenzen der Bewerber/innen mit Schulnoten (1 (sehr gut); 2 (gut), 3 (befriedigend)) bewerten:

Soziale Kompetenzen - Teamfähigkeit; effiziente Kommunikation und Kooperation; interkulturelle Kompetenz. Nachweis etwa durch außerunterrichtliche Aktivitäten (z.B. Mannschaftssport, ehrenamtliches Engagement) oder, falls vorhanden, Kopfnoten.

Kenntnisse über den Beruf - erste Erfahrungen über die Anforderungen im Beruf. Nachweis etwa durch Praktika, Berufsvorbereitungskurse, Nebenjobs oder Hobbys.

Digitale Kompetenzen - Nutzung relevanter Software; Computer- und Programmierfähigkeiten. Nachweis etwa durch im Lebenslauf berichtete IT-Kompetenzen, Informatiknoten, etc.

Analytische Kompetenzen - Reflexionsvermögen; Fähigkeit, nach neuen Informationen zu suchen und diese zu verarbeiten; Fähigkeit, Probleme aus verschiedenen Blickwinkeln anzugehen. Nachweis etwa durch Schul-/Studienleistungen, Hobbys, außerunterrichtliche Aktivitäten (z.B. Schach) etc.

Bitte stellen Sie sich außerdem vor, dass die Bewerber/innen sich ausschließlich in den genannten Attributen unterscheiden, in allen anderen Eigenschaften hingegen nicht. Das bedeutet z.B., dass alle Bewerber/innen über gleich hinreichende Kenntnisse der deutschen Sprache verfügen und in der Nähe Ihres Betriebes wohnen. Die Schulabschlüsse der Bewerber/innen können auch im Ausland erworben worden sein. Sie entsprechen dann den deutschen Schulabschlüssen.

Auf Basis der im Folgenden genannten Eigenschaften bitten wir um Ihre Einschätzung, für wie wahrscheinlich Sie es halten, dass die jeweiligen Bewerber/innen eine Einladung für die nächste Bewerbungsrunde in Ihrem Betrieb erhalten werden.

sehr unwahrscheinlich

1 2 3 4 5 6 7 8 9 10

sehr wahrscheinlich

weiß nicht

C.2 Experiment introduction (German)

Bitte stellen Sie sich einmal unabhängig von der derzeitigen Situation in Ihrem Betrieb vor, dass Ihr Betrieb akut Auszubildende sucht und Ihnen sechs Bewerbungen vorliegen. Ich werde Ihnen die Bewerber einzeln vorstellen und Sie um Ihre Einschätzung bitten, mit welcher Wahrscheinlichkeit Sie diese Personen zu einem Vorstellungsgespräch einladen würden.

Gehen Sie bitte davon aus, dass sich die Bewerber nur in den Merkmalen, Kenntnissen und Kompetenzen unterscheiden, die ich Ihnen jeweils vorlesen werde. In allen anderen Merkmalen gibt es keine Unterschiede.

Bitte schauen Sie auf die Liste

(*INT: Bestätigen lassen, dass Liste Gesprächspartner vorliegt.*)

Dort sind alle Merkmale, Kenntnisse und Kompetenzen aufgeführt, in denen sich die Bewerber unterscheiden. Das wären...

- **Nachname**
- **Geschlecht**
- **Aufenthaltsdauer in Deutschland**
- **Höchster Bildungsabschluss**
- **Soziale Kompetenzen**, d.h. Teamfähigkeit, effiziente Kommunikation und Kooperation sowie inter-kulturelle Kompetenz.
- **Kenntnisse über den Beruf**, d.h. erste Erfahrungen über die Anforderungen im Beruf, z.B. durch Praktika erworben.
- **Digitale Kompetenzen**, d.h. z. B. Nutzung relevanter Software, Computer- und Programmierfähigkeiten.
- **Analytische Kompetenzen**, d.h. z.B. gutes Reflexionsvermögen, die Fähigkeiten neue Informationen zu suchen und diese zu verarbeiten und die Probleme aus verschiedenen Blickwinkeln anzugehen.

Die genannten Kenntnisse und Kompetenzen der Bewerber haben Sie aufgrund der Bewerbungsunterlagen mit den Schulnoten „sehr gut“, „gut“ und „befriedigend“ eingestuft.

C.3 Example task (German)

Bitte stellen Sie sich nun vor, Sie sind akut auf der Suche nach einem oder einer Auszubildenden im Ausbildungsberuf ***[Beruf aus E04a]***.

Ich stelle Ihnen die sechs Bewerber nun einzeln vor und bitte Sie, mir nur aufgrund der genannten Eigenschaften zu sagen, mit welcher Wahrscheinlichkeit Sie die jeweilige Person zu einem Vorstellungsgespräch einladen würden.

Bewerber 1:

Die Person heißt mit Nachnamen Weber, ist männlich, wohnt seit dem Beginn der weiterführenden Schule in Deutschland, hat einen Realschulabschluss, hat sehr gute Kenntnisse über den Beruf, hat sehr gute digitale Kompetenzen, hat sehr gute analytische Kompetenzen, und verfügt über befriedigende soziale Kompetenzen.

Wie wahrscheinlich ist es auf einer Skala von 1 bis 10, dass die folgende Person eine Einladung zu einem Vorstellungsgespräch in Ihrem Betrieb erhält?

1 bedeutet dabei „sehr unwahrscheinlich“, 10 hingegen „sehr wahrscheinlich“.

sehr unwahrscheinlich

1 2 3 4 5 6 7 8 9 10

sehr wahrscheinlich

weiß nicht

References

- Agerström, Jens, Fredrik Björklund, Rickard Carlsson, and Dan-Olof Rooth. "Warm and Competent Hassan = Cold and Incompetent Eric: A Harsh Equation of Real-Life Hiring Discrimination". *Basic and Applied Social Psychology* 34 (2012), 359–366. DOI: [10.1080/01973533.2012.693438](https://doi.org/10.1080/01973533.2012.693438).
- Baert, Stijn, Andrea Albanese, Sofie du Gardein, Jolien Ovaere, and Jarno Stappers. "Does work experience mitigate discrimination?" *Economics Letters* 155 (2017), 35–38. ISSN: 0165-1765. DOI: [10.1016/j.econlet.2017.03.011](https://doi.org/10.1016/j.econlet.2017.03.011).
- Dahl, Malte. "Alike but Different: How Cultural Distinctiveness Shapes Immigrant-Origin Minorities' Access to the Labour Market." *Int. Migration & Integration* 23 (2022), 2269—2287. DOI: [10.1007/s12134-021-00844-y](https://doi.org/10.1007/s12134-021-00844-y).
- Kaas, Leo and Christian Manger. "Ethnic Discrimination in Germany's Labour Market: A Field Experiment". *German Economic Review* 13 (2012), 1–20. DOI: [10.1111/j.1468-0475.2011.00538.x](https://doi.org/10.1111/j.1468-0475.2011.00538.x).
- Oreopoulos, Philip. "Why Do Skilled Immigrants Struggle in the Labour Market? A Field Experiment with Thirteen Thousand Resumes". *American Economic Journal: Economic Policy* 3 (2011), 148–171. DOI: [10.1257/pol.3.4.148](https://doi.org/10.1257/pol.3.4.148).
- Thijssen, L., M. Coenders, and B. Lancee. "Ethnic Discrimination in the Dutch Labor Market: Differences Between Ethnic Minority Groups and the Role of Personal Information About Job Applicants—Evidence from a Field Experiment." *Journal of International Migration and Integration* 22 (2021), 1125–1150. DOI: [10.1007/s12134-020-00795-w](https://doi.org/10.1007/s12134-020-00795-w).
- Thijssen, Lex, Bram Lancee, Susanne Veit, and Ruta Yemane. "Discrimination against Turkish minorities in Germany and the Netherlands: field experimental evidence on the effect of diagnostic information on labour market outcomes". *Journal of Ethnic and Migration Studies* 47 (2021), 1222–1239. DOI: [10.1080/1369183X.2019.1622793](https://doi.org/10.1080/1369183X.2019.1622793).