Examining discrimination in asynchronous video interviews: Does cultural distance based on country-of-origin matter?

Rene Arsenault, *University of Laval*

Nicolas Roulin, *Saint Mary's University*

Abstract

We conducted two studies to investigate how cultural differences based on country of origin influence the selection process in an asynchronous video interview (AVI) context. We drew upon the GLOBE cultural value dimensions and individual measures of prejudice to examine if raters evaluate job applicants who are more culturally-dissimilar to them more negatively than culturally-similar applicants. Professionals with hiring experience from the U.K. were recruited via the Prolific platform and asked to watch and evaluate pre-recorded video responses from five culturally diverse applicants. Results across both studies were only somewhat consistent with the GLOBE framework. For instance, raters did demonstrate a strong preference for Canadian and South African interviewees over other countries. Right-wing authoritarianism and social dominance orientation were non-significant in moderating how evaluations were assigned; however, ethnocentrism levels did modestly impact evaluations in Study 2. This research is the first to investigate how cultural factors can impact the selection process in an AVI context. As the number of organizations that rely on virtual interviews increase and globalization makes it likely for applicants and interviewers to be from different cultural backgrounds, our research is highly relevant in understanding the impact of these elements on hiring decisions.

Keywords: cross-cultural selection, asynchronous video interviews, GLOBE framework, COVID pandemic, recruitment

Introduction

The COVID-19 pandemic has created unique challenges for recruitment and selection efforts. For many organizations, face-to-face (FTF) interviews are no longer the only option. Alternative mediums such as using video-conference software (e.g., Zoom, Teams, Skype) and asynchronous video interview platforms (AVIs - e.g., HireVue, ModernHire, Aon VidAssess) for interviewing job applicants are increasingly used in the selection process (Griswold et al., 2021; Maurer & Maurer, 2020). As an example, the two biggest AVI providers, Modern Hire and HireVue, claim to work with over half and a third of the Fortune 100 companies respectively. And, HireVue recently reported having hosted over one million AVIs in September of 2021 alone (HireVue, 2021). The implications of transitioning to and relying largely (or solely) on virtual platforms for organisational selection processes merits investigation (Carnevale & Hatak, 2020). Further, there is a growing recognition of the value of workforce diversity (Scullion et al., 2016). This recognition reflects the needs of a globalized economy which requires organizations to recruit employees from diverse cultural backgrounds with strong cultural awareness, language competencies and varying perspectives (Zheng & Menzies, 2015). To stay competitive, multinational corporations need to strategically recruit, select, and deploy their global talents.

Traditional FTF interviews are highly beneficial in gathering additional job-relevant information beyond a resume about candidates during the selection process, in order to make judgments of employment suitability (Huffcutt & Arthur, 1994; Schmidt & Hunter, 1998). In addition, structured interview formats are relatively resistant to many types of biases, such a similarity between the applicant and interviewer in terms of gender or race/ethnicity (Kith et al., 2022; Levashina et al., 2014; McCarthy et al., 2010; Pogrebtsova et al., 2020). Technology-mediated interviews like AVIs can offer improved efficiency for the initial screening process,

particularly for positions with numerous applicants and/or with geographic challenges, such as when hiring out of region/country. AVIs are also quite standardized, and thus could benefit from the same protection against biases as structured FTF interviews. However, very little is known about the effectiveness of using AVIs in the selection process, and their potential for biased selection decisions (Lukacik et al., 2022). Findings from in-person interviews do not necessarily translate to video-conference interviews (Blacksmith et al., 2016). They might arguably also not translate to AVIs, for instance because of their restrictive nature (i.e. one-way communication).

Interviewers often form first impressions of applicants in the first few seconds of the interview, and engage in cognitive processes to confirm these initial impressions (Derous et al., 2016). Such process could take place both in FTF interviews and AVIs. For example, visible elements like race, ethnicity, or non-native accents could also influence performance evaluations in AVIs. In addition, technologically-related elements (i.e. background in video-recordings, internet connection and audio quality) may trigger bias unique to AVIs (Lukacik et al., 2022; Roulin et al., 2022). Such biases could prevent organizations from reaching diversity objectives or subject them to discrimination lawsuits. For example, the U.S. (Title VII of the Civil Rights Act of 1964), Canada (Canadian Charter of Human Rights), the U.K (Equality Act of 2010), or the EU (Article 21 of the Charter of Fundamental Rights) have strict laws against various forms of hiring discrimination. Particularly when hiring across national borders or in a region with a multi-cultural workforce, organizations must ensure that their selection tools and procedures (including emerging technologies such as AVIs) are not discriminating applicants based on their country of origin or nationality.

The present research contributes to the literature on cross-cultural personnel selection in several important ways. First, it examines whether findings about how biased decisions based on

country of origin or nationality using traditional selection methods (Derous et al., 2012; Petersen & Dietz, 2005; Veit & Thijsen, 2021) also take place in an AVI context. Such an investigation is particularly significant given that the COVID-19 pandemic has forced many organizations to interview applicants primarily through online mediums (Maurer & Maurer, 2020). Second, extant research has generally compared interviewees from in-group (e.g., local applicants) vs. out-group (e.g., one type of immigrant). In contrast, we explore multiple out-group applicants that vary in terms of their cultural similarity to the interviewer, drawing upon GLOBE cultural dimensions (House et al., 2004). Namely, in two complementary studies, British professionals with hiring experience were asked to evaluate video-recordings from five applicants from countries of varying degrees of cultural similarity. As such, our study is one of the firsts to examine whether bias against out-group members are stronger for applicants from very culturally-distant countries (e.g., India or South Africa for British raters) than from more culturally-similar countries (e.g., Canada). Finally, building on theories of interview evaluations (Huffcutt et al., 2011) and the motivational model of ideology and prejudice (Duckitt & Sibley, 2017), we explore whether interviewer-level individual differences (e.g., ethnocentrism) can moderate the impact of cultural distance on evaluations.

Asynchronous Video Interviews

AVIs, also known as digital interviews, are conducted without live interaction. Typically, candidates are invited to read questions pre-selected by the hiring organizations and video-record their answers (Brenner et al., 2016). AVIs can help increase capacity for handling a large number of recruits, streamline selection procedures, decrease administrative strain by ensuring only the most qualified candidates access the next step of the selection process, and increase organizational efficiency for HR teams storing and sharing applicant videos throughout the

selection process (Mejia & Torres, 2018; Stone et al., 2015). However, AVIs also present several weaknesses that organizations must be aware of. For example, a meta-analysis by Blacksmith et al. (2016) found that interviewees are generally more skeptical and less accepting of technology-mediated interviews compared to FTF interviews. Yet, this meta-analysis was based almost exclusively on videoconference and telephone interviews, and it is therefore possible that AVIs require separate consideration.

In addition, the literature on AVIs is still very young. To date, most AVI research has focused on the applicant side, for instance examining applicants' reaction to and behaviors in AVIs (Basch et al., 2021; Hiemstra et al., 2019; Langer et al., 2018). Research focused on the interviewer/rater side of AVIs is scarce, and limited to general reactions about the practicality or usefulness of the technology (e.g. Basch & Melchers, 2021; Mejia & Torres, 2018). Only two studies have examined biases, but focused on aesthetical features like attractiveness (Torres & Gregory, 2018) or video background (Roulin et al., 2022). Research examining how cross-cultural differences could impact evaluations in AVIs is lacking, and warrants investigation given the increase in AVI use in practice (and accelerated by the COVID pandemic) and its value in facilitating the selection of a broader (e.g., international or cross-cultural) population of job applicants.

Further, the increasingly transactional nature of electronic hiring systems may also lead to negative perceptions, such as an impersonal feeling, a lack of real-time feedback and a feeling of not being able to portray oneself in a realistic manner (Guchait et al., 2014). Basch et al. (2020) suggest that AVIs are less accepted than traditional FTF interviews because they are seen as less fair. Whether or not AVIs are fair and an unbiased tool in the selection process has relevance to organizations because fairness perceptions are positively associated with perceived

organizational attractiveness (e.g., Walker et al., 2015). Finally, there are also privacy concerns related to using new technologies for selection procedures (Stone-Romero et al., 2003). Indeed, AVIs are perceived to be "creepier", less personal, and associated with more privacy concerns than video-conference interviews (Langer et al., 2017). In sum, AVIs requires separate investigation(s) from traditional FTF interview research to understand the implications and consequences of their use in the selection process (Lukacik et al., 2020).

Given that AVIs are visual recordings, they can still present evaluators with similar surface level demographic information (i.e., about race or ethnicity) and audio cues (i.e. non-native accents) found in FTF interviews that may suggest the applicant is from a foreign country. Therefore, is it possible that AVIs create opportunities for discrimination and biased selection decisions. However, to date no research has examined the presence of culturally-based biases in AVIs. A large body of literature over the past 50 years has investigated ethnic or racial hiring discrimination and consistently found it to be a real issue in FTF interviews (Zschirnt & Ruedin, 2016). However, much less research exists examining hiring discrimination/biases based on country of origin/nationality. Therefore, the present study in interested in investigating whether AVIs may be biased based on an applicant's country of origin/nationality.

Bias in Selection

As mentioned above, research exploring cultural biases exclusively in an AVI context is lacking. We therefore draw on the FTF interview literature for insights. Traditionally hiring discrimination literature has focused on differences between Black and White applicants in the U.S. (e.g., Wexley & Nemeroff, 1973). And, a meta-analysis of 28 studies of U.S. labor markets found no evidence of change in the level of hiring discrimination against Black applicants over the past 25 years, and only modest evidence of a decline in discrimination against Latino

applicants (Quillian et al., 2017). Globally, a meta-analysis of 738 correspondence tests from 43 studies conducted between 1990 and 2015 found that ethnic and racial discrimination has remained widespread across OECD countries (Zschirnt & Ruedin, 2016). Beyond race, research has also examined the impact of native versus non-native accents on hiring decisions. For example, Segrest-Purkiss et al. (2006) found that applicants with an ethnic name were viewed less positively by interviewers when speaking with accents (vs. without accents). Accents can also influence perceptions of intelligence, kindness, social status, economic class, national origin, and ethnicity (e.g., Lippi-Green, 2012; Nesdale & Rooney, 1990). In the U.S., French accents often are associated with sophistication, whereas Asian accents are associated with high economic and educational attainments (Cargile, 2000; Lippi-Green, 2012). And most recently, Hideg et al., (2022) conducted an interdisciplinary review and research agenda for non-native accents in the workplace, highlighting the importance of distinguishing between non-native accents that impact comprehensibility, versus those non-native accents that do not.

Most of the findings discussed above can be explained by psychosocial processes described in Tajfel and Turner's (1986) social identity theory and its related cousin self-categorization theory (Turner & Oakes, 1986). Such processes lead to in-group favoritism where people give preferential treatment (i.e., better evaluations) to others when they are perceived to belong to the same in-group. Applicants who are from the same race as hiring managers (in-group) receive more favorable interview assessments and are more likely to receive job offers than are applicants who are from a different racial group (out-group; see Dovidio & Gaertner, 2000). However, more recent work has started to go beyond an in- vs. out-group approach and compared different group of applicants, for instance based on cultural similarity levels.

The Role of Culture in Biased Selection Decisions

Hiring discrimination can depend on the level of cultural (dis)similarity between the applicant and the hiring manager. For instance, Veit and Thijsen (2021) compared reactions of employers in Western European countries to job applications from majority applicants, minority applicants culturally-similar to them (domestic-born and/or European origin), or minority applicants culturally-dissimilar or distant to them (Middle Eastern or African origin). They found especially high levels of discrimination for the latter group. Similarly, strong evidence of hiring discrimination in Western countries against Arabs or Muslims (i.e., a culturally-distant group) was also found in a recent meta-analysis (Bartkoski et al., 2018). More generally, another meta-analysis of 97 field experiments conducted across nine countries in Europe and North America found that hiring discrimination fell on a continuum ranging from non-White immigrants (high levels of discrimination) to non-White native applicants (moderate) to White native applicants (lowest; Quillian et al., 2019).

The personnel selection literature offers valuable insights as to why cultural distance between the applicant and interviewer might bias selection decisions. First, the objectives of a typical interviewer commonly include determining how well the applicant's qualifications match the job requirements (Person-Job fit) and how their personality, goals, or values match the culture of the hiring organization (Person-Organization fit; Kristof-Brown et al., 2002). However, using person-organization fit criteria to 'qualify' a candidate leads to unavoidable biases, for example, because interviewers' fit assessments often involve shortcuts like selecting the candidates that more closely resemble them (Cable & Judge, 1997). This increases the risk that interviewers negatively evaluate applicants who are different from them in terms of demographic, personality or attitudinal characteristics, and triggers similar-to-me biases (Sears &

Rowe, 2003). Such biases are likely to be reduced with more structured interviews (Levashina et al., 2014; McCarthy et al., 2010).

   In the context of cross-cultural hiring with AVIs, this suggests that interviewers are likely to evaluate applicants from a culture similar to theirs more positively that those from a culture dissimilar to theirs. This argument is aligned with several models of interviewee performance. For instance, Huffcutt et al., (2011) suggested that the relationship between interviewees' objective interview performance and how they are evaluated by interviewers can be impacted by differences in cultural backgrounds between the two actors. They argued (p.361) that "interviewers may see candidates from a different culture as having a lesser fit with their company even though their performance on the job might not be impacted." Similarly, Manroop, et al.'s (2013) proposed a model of how cross-cultural differences impact interview outcomes, and argued that "interviewers are more likely to make negative judgment about the job candidates who respond to questions contrary to cultural expectations than candidates who respond to questions according to cultural expectations" (p. 3522). Both points imply that cultural differences in behavioral expectations could impact interviewee evaluations. For example, according to GLOBE research, British culture values of assertiveness are lower than Indian cultural values of assertiveness. Indian applicants may assert themselves in a manner that is considered ideal in their own culture, but cocky or arrogant to a British interviewer (i.e., perception of excessive assertiveness). Similarly, lower levels of power distance values for British interviewees as compared to South African values in this dimension may create expectations for less formal/hierarchical interactions. Based on the discussion above, we predict:

Hypothesis 1: Interviewers assign higher evaluations to applicants that are culturally similar to them, and lower evaluations to applicants that are culturally dissimilar in an AVI context.

We propose to rely on the GLOBE model (House et al., 2004), to estimate the level of cultural distance or dissimilarities between interviewers and interviewees and test this hypothesis. Nearly a decade ago, in the Oxford handbook of personnel assessment and selection, Steiner (2012: 741) mentions how '*few studies on personnel selection internationally have systematically studied cultural variables associated with their effective application and few cultural variables are represented in the studies conducted*', indicating the lack of and need for more research within this area. Although other iconic cultural frameworks (i.e., Hofstede, 1980; Schwartz, 2006) overlap with GLOBE in terms of cultural dimensions (i.e., power distance, collectivism, gender egalitarianism) and regional 'groupings' (i.e., Anglo, Confucian Asia) we see the GLOBE framework as more comprehensive and less subjected to criticisms in the literature. For example, the GLOBE framework offers nine cultural dimensions distinguishing between values and practices whereas Schwartz's (2006) value orientations cover only seven dimensions, and Hofstede's (1980) six. The GLOBE framework also presents the highest number of regional groupings at 10, whereas Hofstede's (1980) framework consisted of seven cultural groupings, and Schwartz's (2006) framework six. Noteworthy criticisms in the literature surrounding Hofstede's work (see Brewer & Venaik, 2011; Javidan et al., 2006; McSweeney, 2002) also informed our decision to use the GLOBE framework. There also exists some relevant personnel selection research that has used the GLOBE model empirically (Allen & Vardaman, 2017; Fell & König, 2016) and theoretically (Arseneault & Roulin, 2021; Farndale & Sanders, 2017).

Potential Interviewer-level Moderators

In addition to the general prediction that the larger the cultural distance between interviewers and applicants, the lower the performance evaluation applicants will receive, we

also believe there could be some boundary conditions to this general effect. In their model, Huffcutt et al. (2011) argued that there are individual differences in how interviewers process information from the interviewee, and emphasized that some interviewers are more prone to bias than others. Similarly, Derous et al.'s (2016) model of interviewer bias also highlight the moderating role played by interviewer characteristics. In the context of cross-cultural AVIs, we propose to examine three types of interviewer individual differences: ethnocentrism, social dominance orientation, and right-wing authoritarianism.

Ethnocentrism.

Ethnocentrism refers to levels of openness (or lack thereof) to foreign cultures or outsiders (Neuliep & McCroskey, 1997). Those high on ethnocentrism view their culture as the 'center' of the world, superior to, and a role model for other cultures. They also generally dislike interacting with members of foreign cultures. Levels of ethnocentrism vary across countries (Neuliep et al., 2001). People higher on ethnocentrism are less willing to communicate with individuals from other cultures than members of their own cultures (Lin et al., 2005) and, particularly relevant to the present study, tend to evaluate job applicant from other cultures (Neuliep et al., 2005) or with non-native accents (Neuliep & Speten-Hansen, 2013) more negatively. As such, we propose that ethnocentrism may play an important role in moderating how (British) raters evaluate applicants from different cultures. More precisely, those high on ethnocentrism should be especially likely to assign negative evaluations to applicants from culturally-distant countries (i.e., with non-native accents, demographic and culturally foreign behaviors), whereas those low on ethnocentrism should be more open to (and thus less impacted by) such cultural differences:

Hypothesis 2: The impact of the cultural distance between the applicant and the interviewer on performance evaluation in an AVI context is moderated by interviewers' level of ethnocentrism, such that the relationship is stronger when ethnocentrism is higher.

Social Dominance Orientation and Right-Wing Authoritarianism.

Next, we consider how two major social attitudinal predictors of prejudice, right wing authoritarianism (RWA) and social dominance orientation (SDO), could impact evaluations interviewers assign to applicants from various cultural backgrounds. Duckitt and Sibley (2017) discuss the complimentary effects of RWA and SDO in their dual-process model of prejudice. Their model explains how the underlying personality dimensions of RWA (i.e., low agreeableness and high conscientiousness) and SDO (i.e., low agreeableness) represent social and psychological bases of personality that contribute to dangerous and competitive worldview beliefs (i.e., embeddedness values) that lead to an in-group preference for order, structure, stability, and security. A large body of research has found RWA and SDO to be powerful predictors of prejudice (Sibley & Duckitt, 2008), and prejudice does moderate interviewers' reactions to out-group job applicants (Petersen & Dietz, 2005).

SDO is an individual difference variable that indicates support for the "domination of 'inferior' groups by 'superior' groups" (Sidanius & Pratto, 1999), p. 48). SDO is a valid predictor of a range of biases across cultures (Lee et al., 2011). This is in part because SDO is highly correlated with hierarchy (Duckitt & Sibley, 2017), and thus high-SDO people believe that there are (and should be) status differences among social groups (Pratto et al., 1994). SDO also plays a role in hiring decisions. For example, high-SDO Americans are particularly unlikely to select a potential team member from a low-status group (Umphress et al., 2008) or to recommend hiring Asian or Latino applicants (Hansen & Dovidio, 2016). As such, we predict

that interviewers' level of SDO will also affect the strength of the relationship between cultural distance and interviewers' evaluation of applicants. Namely, high levels of SDO will amplify the negative evaluations resulting from large cultural distances between interviewers and interviewees whereas low levels of SDO will weaken this effect:

Hypothesis 3: The impact of the cultural distance between the applicant and the interviewer on performance evaluation in an AVI context is moderated by interviewers' level of social dominance orientation, such that the relationship is stronger when social dominance orientation is higher.

Finally, RWA relates to the desires to protect and enhance the self and the in-group leading to greater in-group favoritism, and in some cases prejudice towards out-groups (Altemeyer, 1988). Therefore, those who score high on RWA are more likely to favor in-group versus out-group members. For instance, Charles-Toussaint and Crowson (2010) found that RWA positively correlated with American students' prejudice towards international students. The role of RWA has also been empirically investigated in the employment discrimination literature. For example, Petersen and Dietz (2000) found that only high-RWA West-German raters (in-group) discriminated against East-Germans applicants (out-group) when instructed to do so. Therefore, we predict that interviewers' level of RWA moderates the relationship between cultural distance and evaluation of applicants. Higher levels of RWA should amplify the negative evaluations resulting from a large cultural distance between interviewers and interviewees, whereas lower levels of RWA should weaken it.

Hypothesis 4: The impact of the cultural distance between the applicant and the interviewer on performance evaluation in an AVI context is moderated by interviewers' level of right-wing

authoritarianism, such that the relationship is stronger when right-wing authoritarianism is

higher.

Study 1

Study 1 represents an initial attempt to test our hypotheses using a sample of British

raters with hiring experience, evaluating applicants from five countries. Table 1 shows the

GLOBE's nine cultural dimensions value scores for the six countries relevant for Study 1, as

well as a composite distance score highlighting cultural differences between the country of origin

of our raters (the U.K.) and the five countries our interviewees were from (Canada, India,

Poland, Spain, South Africa). According to our main hypothesis, a British interviewer should

assign higher scores to applicants from countries with more similar cultures (e.g., Canada) and

lower scores to applicants from more dissimilar ones (e.g., South Africa).

Table 1 – Country comparison of GLOBE cultural dimensions

	U.K.	Canada	Poland	Spain	India	South Africa
Performance orientation	**5.90**	6.15	6.12	5.80	6.05	4.92
Assertiveness	**3.70**	4.15	3.90	4.00	4.76	3.82
Future orientation	**5.06**	5.35	5.20	5.63	5.60	5.20
Humane orientation	**5.43**	5.64	5.30	5.69	5.28	5.07
Institutional collectivism	**4.31**	4.17	4.22	5.20	4.71	4.30
In-group collectivism	**5.55**	5.97	5.74	5.79	5.32	4.99
Gender egalitarianism	**5.17**	5.11	4.52	4.82	4.51	4.26
Power distance	**2.80**	2.70	3.12	2.26	2.64	3.65
Uncertainty avoidance	**4.11**	3.75	4.71	4.76	4.73	4.79
Total distance score		**2.28**	**2.54**	**3.90**	**3.97**	**4.61**

Note. GLOBE value scores

Methods

Sample

We recruited a total of 100 British participants with HR/interview-related experience

through the Prolific online recruitment platform. The mean age of participants was 44.5 (*SD* =

12.05). Participants were 53% female and 90% White (3% Black, 5% Asian, 2% Other). They had an average of 8.3 years of HR-related experience ($SD = 12.05$), 19% had experience using AVI interviews, and on average had lived 1.1 years abroad (i.e., outside of U.K. - $SD = .35$). Sensitivity analyses (using G-Power) suggested our sample of N=100 was sufficient to detect effect sizes as small as f = .14 (i.e., an eta-squared of .02), when examining Hypothesis 1 (using a repeated-measure ANOVA - with 5 groups/countries). When examining Hypotheses 2-4 (repeated-measure ANCOVAs with a continuous covariate), our analysis suggests that our sample allowed us to detect effects as small as $f^2 = .13$ (i.e., eta-squared = .12). See our online supplement for more detail.

Procedure

Participants were invited to view and score the video-recorded responses of five applicants. Participants were instructed to act as interviewers in charge of hiring for a management associate position in a multinational bank. Before watching the interview videos, participants were provided with a job description, the four interview questions applicants answered, and a scoring rubric (see online supplement). Each participant was then asked to watch and evaluate four video-recorded responses from five interviewees (i.e., a total of 20 videos). Each applicant was from a different country (i.e., Canada, India, Poland, South Africa, and Spain - presented in a random order). After watching the fourth video for each applicant, participants were asked to complete/evaluate them on several job-related constructs. After watching all 20 videos and completing evaluations, participants were asked to also complete several measures (i.e., Ethnocentrism, SDO, RWA) and answer demographic questions.

The interviewees' videos were mock interviews conducted as part of a separate study by the authors (with over 50 interviewees for each of the five countries). We selected a total of 15

interviewees (i.e., three per country) who demonstrated equivalent performance on the four

interview questions, as rated by trained coders using behaviorally-anchored rating scales (BARS

– see online supplement). All coders were undergraduate students in psychology from different

cultural background (but not citizens of the countries used in the study), who received an initial

one-hour training by one of the co-authors (including getting familiar with the rating scales,

coding a set of videos together, comparing score, discussing discrepancies). Coders were then

assigned a set of videos to independently score, as well as 10 videos that all coders scores to

confirm intercoder agreement (with ICCs ranging from .79 to .95 for the four questions). This

allowed us to ensure that differences in ratings by participants in this study were caused by

biased assessments (i.e., due to cultural differences) and not because of the objective quality of

the interviewee's responses. Interviewee demographics for each 'triad' were balanced with

respect to gender (i.e., 1-2 males, 1-2 females), age (i.e., mixed ages in 20's and 30's), and

ethnicity was controlled (i.e., all South Africans were Black, Indians East Asian, and remaining 9

interviewees all White/Caucasian). Videos length was also largely consistent with reported

maximum response time allocated to applicants, and actual use of that time by applicants, in

practice (see Dunlop et al., 2022). Our participants were randomly assigned to watch one of the

three interviewees from each country. This approach (vs. using only one interviewee per

country) also guarantees that effects are due to cultural differences and not unique characteristics

of a specific individual, thus increasing external validity.

Measures

   Interview performance ratings. Participants used a rating scale developed by Gorman et

al. (2018) designed to evaluate applicants in AVIs. We used six of the original twelve

constructs/work-related attributes including general intelligence, conscientiousness, interpersonal

skills, or leadership, after removing constructs (e.g., education and training, creativity, job knowledge) that were less directly applicable to our study context (see online supplement for final version of the six constructs used in this study). Participants used a 5-point Likert scale to evaluate applicants on each construct. We combined ratings of the six constructs into an overall interview evaluation score for each of the five countries (α = .85 - .92) for our analyses.

Ethnocentrism. We used an ethnocentrism scale developed by Neuliep and McCroskey (1997) to measure participants' attitudes towards foreigners. This version includes 22 items using a 5-point (strongly disagree to strongly agree) Likert scale (α = .90). Six of the 22 items are reverse coded. An example item is *"Most people from other cultures just don't know what's good for them"*.

Social Dominance Orientation. We used the SDO7 developed by Ho et al. (2015) to measure participants' levels of social dominance orientation. This version includes 16 items (α = .94) grouped into 4-dimensions, Pro-trait dominance, Con-trait dominance, Pro-trait anti-egalitarianism and Con-trait egalitarianism. A 7-point (strongly favor – strongly oppose) was used. The Con-trait items are reverse-scored before computing the composite scale mean. An example item is "*Some groups of people must be kept in their place.*"

Right-wing Authoritarianism. We used the Very Short Authoritarianism (VSA) scale developed by Bizumic and Duckitt (2018) to measure participants' levels of authoritarianism. It includes six items (α = .81) graded on a 5-point (strongly agree – strongly disagree) Likert scale, which capture the three content dimensions of Altemeyer's widely used Right Wing Authoritarianism (RWA) scale. An example item is "*God's laws about abortion, pornography, and marriage must be strictly followed before it is too late*".

Results

Preliminary analyses revealed some missing values, and two potential univariate outliers (SDO, $z = 3.35$; Ethnocentrism, $z = 4.19$), but no severe item-response related issues. Removing those two individuals did not influence our results. Additionally, all participants passed the three attention checks embedded within our measures. Therefore, we decided to keep all 100 data points for our main analyses.[1] Table 2 summarizes correlations among study variables for our study. As expected, all three measures of prejudice (i.e., RWA, SDO, Ethnocentrism) significantly correlated with each other. Age significantly correlated with HR experience ($r = .53$, $p < .01$). None of our measures (SDO, RWA, Ethnocentrism) significantly correlated with country evaluation scores.

Table 2 – Correlations among study variables

		M	SD	1	2	3	4	5	6	7	8	9
1	Age	44.52	12.05									
2	HR Exp.	8.24	8.20	.53**								
3	RWA	2.64	0.75	-.10	.12							
4	SDO	1.96	0.69	.08	.01	.45**						
5	ETHNO	2.00	0.51	-.04	-.15	.42**	.53**					
6	Poland	2.94	0.66	-.04	-.18	.08	.02	.03				
7	Spain	2.78	0.83	-.13	-.03	.10	-.11	-.09	.45**			
8	India	3.41	0.79	-.17	-.11	.11	.01	-.09	.29**	.36**		
9	South Africa	3.55	0.74	.01	-.15	-.02	.01	-.03	.08	.07	.20*	
10	Canada	3.57	0.71	.04	-.04	.15	-.12	.04	.09	.04	.21*	.04

Note. $* p < .05$, $** p < .01$. *Poland/Spain/India/South Africa/Canada variables refer to evaluation scores received by interviewee in those respective countries. HR Exp.= years of experience in an HR role, RWA = right-wing authoritarianism, SDO = social dominance orientation, ETHNO = ethnocentrism

Cultural Differences

---

[1] We also replicated all our analyses (in both studies) after removing non-White participants. Results were equivalent to those with the full sample (although the effects were slightly weaker for some moderations).

As noted in Table 1, we computed a total distance score for each of the five countries (vs. the U.K.) from the GLOBE scores, which resulted in the following predicted order, from most similar to dissimilar: Canada (2.28), Poland (2.54), Spain (3.90), India (3.97) and South Africa (4.61). To test Hypothesis 1, we then conducted a within-subjects repeated measure ANOVA to examine whether mean score differences across our five countries were statistically significant and aligned with the order predicted above. There was a statistically significant difference in mean scores across the five countries, $F(4, 396) = 29.394$, $p < .001$, $\eta^2 = .229$.. The left part of Table 3 summarizes descriptive statistics for the five countries. Pairwise comparisons further showed that evaluations of applicants from Canada ($M = 3.57$, $SD = 0.71$), South Africa ($M = 3.55$, $SD = 0.74$) and India ($M = 3.41$, $SD = 0.79$), were not statistically different from one another (i.e., all $p > .09$). However, evaluations of applicants from these three countries were significantly higher than evaluations of applicants from both Poland ($M = 2.94$, $SD = 0.66$; all $p < .001$) and Spain ($M = 2.78$, $SD = 0.83$; all $p < .001$). Overall, these results are only partly consistent with our first hypothesis that country level evaluations would decrease as cultural distance from U.K. increases. While the results for Canada, Poland, and Spain were generally aligned with our predictions, South African and Indian participants (most culturally-distant from the U.K.) were rated significantly higher than expected.

Moderation Analyses

To test for the moderating role of RWA, SDO and ethnocentrism (Hypotheses 2-4), we standardized them and included them separately as covariates in repeated-measure ANCOVAs. Results are presented in Table 3. For ethnocentrism, and contrary to Hypothesis 2, the country of origin x ethnocentrism interaction was not significant. For SDO, and contrary to Hypothesis 3, the country-of-origin x SDO interaction was also not significant. Finally, similar results were

found for RWA, with a non-significant country-of-origin x RWA interaction, thus not supporting

Hypothesis 4. All three interaction effects remained non-significant in additional ANOVAs

incorporating respondents' experience in human resources as an additional covariate.

Table 3. Repeated-measure ANCOVA results for Study 1

	F	p	η^2	η_p^2
Country-of-origin	29.241	$< .001$	.229	.230
Ethnocentrism	0.312	.578	.001	.003
Country x ethnocentrism	0.483	.721	.004	.005
Country-of-origin	29.277	$< .001$	.229	.230
Social Dominance Orientation	0.443	.507	.001	.004
Country x SDO	0.605	.646	.005	.006
Country-of-origin	29.224	$< .001$	.229	.230
Right-Wing Authoritarianism	2.030	1.57	.004	.020
Country x RWA	0.427	.774	.003	.004

Note. F-values are based on the Greenhouse-Geisser correction

Robustness Checks

To further confirm our findings, we also explored how evaluations provided by UK raters

in Study 1 for the interviewees from the five countries compared to objective performance scores

(i.e., BARS scores from trained raters obtained in a separate study). As illustrated in the right

part of Table 4, interviewees from India, Spain, and Poland received evaluations that were

significantly lower than the objective BARS scores, which provides further evidence of

culturally-biased evaluations. But interviewees from Canada and South Africa did not.

Table 4 – Mean scores in performance/evaluation across five countries

Country	Participants ratings *M (SD)*	Confidence Intervals	BARS *M (SD)*	Mean Difference	*t*-test values	*p*-value	Cohen's *d*
Canada	3.57 (0.71)	3.66 – 3.90	3.75 (0.97)	0.18	$t(99) = 1.497$	.136	.212
South Africa	3.55 (0.74)	3.40 – 3.70	3.67 (0.65)	0.12	$t(99) = 1.218$	.225	.172
India	3.41 (0.79)	3.26 – 3.57	3.67 (0.65)	0.26	$t(99) = 2.541$	.012	.359
Poland	2.94 (0.66)	2.81 – 3.07	3.67 (1.00)	0.73	$t(99) = 6.092$	.001	.862
Spain	2.78 (0.83)	2.62 – 2.95	3.58 (0.67)	0.80	$t(99) = 7.499$	.001	1.06

We also replicated our analyses while controlling for participants' level of hiring experience (operationalized as number of years working in human resources), to see whether biased evaluations against applicants from different cultures would be weaker (or stronger) for more experienced respondents. Our within-subjects repeated measure ANCOVA using HR experience as a covariate did not reveal any significant interaction between country of origin and experience, $F(4,388) = 0.347$, $p = .831$). But the main effect of the country of origin remained significant, $F(4, 388) = 15.089$, $p < .001$.

Additionally, we explored whether some of the unexpected results (e.g., higher-than-anticipated ratings for South African interviewees) could be caused by unusually-high ratings obtained by one of the three applicants in each country. We thus compared the ratings obtained by the three interviewees in our five countries using ANOVAs. We found no difference between the three Canadian, $F(2,97)=2.361$, $p = .100$, Polish, $F(2,97)=1.370$, $p = .259$, or Spanish, $F(2,97)=2.390$, $p = .097$, interviewees. We found significant differences between the three Indian, $F(2,97)=3.955$, $p = .001$, and South African, $F(2,97)=7.284$, $p = .001$, interviewees. Yet, in both cases, paired comparisons showed that one interviewee scored significantly *lower* than the other two (who obtained equivalent scores), thus alleviating potential concerns that one high performing candidate 'raised' evaluation scores for the entire group.

Discussion

Study 1 findings showed that evaluations of the Canadian, South African, and Indian interviewees were significantly higher that evaluations of Spanish and Polish interviewees. These scores and ranking were only partially aligned with our initial predictions based on differences in

GLOBE value scores between the U.K. and these five countries. Past research has generally focused on comparing applicants from a single in-group vs. out-group design, such as one racial or ethnic minorities group evaluating another. For example, Veit and Thijsen (2021) found that Western European employers discriminated against applicants from Middle Eastern / African origins. In contrast, we examined potential discrimination against multiple out-groups, which differ in terms of their cultural distance from the raters' country of origin (the U.K.). Yet we only found clear evidence for discrimination against Spanish and Polish (i.e., Caucasian/White European) applicants. This suggests that other factors may have influenced how participants evaluated interviewees and represent boundary conditions to the effect of culture on interviewers' evaluations. We discuss a few potential alternative explanations below, and then introduce Study 2, which was designed to further test them (alongside most of our initial hypotheses).

One possible explanation for our findings is that language barriers and skills influenced evaluation scores assigned to interviewees. Canada, India, and South Africa all have English as one of their official national languages, whereas Poland and Spain do not. Research examining the role accents play in impacting interview evaluations has mixed findings and is worth discussing within the context of the present study (Hideg et al., 2022). For example, accents have been found to both positively (Deprez-Sims & Morris, 2010) and negatively impact interview evaluations (Campbell-Kibler, 2007; Lippi-Green, 2012; Segrest Purkiss et al., 2006), but more research is needed to clarify the context(s) influencing the direction of such effects (Deprez-Sims & Morris, 2013). Important considerations such as position type (see Hosoda et al., 2012) and whether the job is customer facing versus non-customer facing (Timming, 2017) may also explain whether accents significantly impact evaluations. It is possible that accents in the present

study significantly and negatively impacted how evaluations were assigned to some groups (i.e., Polish, Spanish), but research suggests that when English proficiency is strong enough to communicate effectively, accents are less likely to negatively impact evaluations (Creese & Kambere, 2003; Hideg et al., 2022). We therefore argue that this first explanation is unlikely, since we carefully selected interviewees with similar English language proficiency levels (except native English-speaking Canadians) for Study 1. Of course, interviewees from all countries in our Study 1 had noticeable accents, which were likely perceived by raters. And, the British/English root accents have a common ancestor that creates familiarity and general fluency, which could have favored the Canadian, South-African, or Indian applicants as compared to the Polish or Spanish ones. Yet, we noticed that Indian applicants' accent was noticeably stronger (i.e., at times slightly more difficult to comprehend) than were accents observed from Polish and Spanish applicants. Therefore, the accent might have helped the South-African interviewees obtain higher-than-expected evaluations, but likely not the Indian interviewees. Therefore, in Study 2, we took additional precautions to further limit differences in English skills and accents across interviewees from different countries.

In addition, we initially selected interviewees across the five countries similar levels of objective interview performance (from BARS scores from trained raters, ensuring that all applicants provided responses of similar quality). Additional analyses comparing the evaluations of UK interviewers to objective performance suggest that Spanish and Polish interviewees received particularly lower-than-objectively-expected evaluations, whereas South-African interviewees did not (and Indian interviewees received only slightly-lower scores). It is therefore possible that other unmeasured differences between our interviewees (beyond culture) impacted Study 1 results. Thus, in Study 2, we took extra precautions to limit differences between

interviewees for three factors that have been shown to impact hiring decisions: perceived warmth and competence (e.g., Fiske et al., 2002; Krings et al., 2011), as well as perceived attractiveness (e.g., Tews et al., 2009).

A final potential explanation for the main findings from Study 1 is that U.K raters might have considered their shared history with Canada, South Africa, India, that is, the three countries with British roots. This might explain why raters gave applicants from those countries higher scores than those from Poland and Spain. The employment interview is a context where cultural factors have been found to influence applicant behaviors (e.g., Sandal et al., 2014), and interviewer preferences for culturally similar behaviors (e.g., Derous, 2017). And, findings for some of our countries confirm that cultural similarities/dissimilarities (i.e., from GLOBE cultural dimensions) influence applicant evaluations in an AVI context too, for instance with Canadian interviewees being rated higher than Spanish or Polish ones. Yet, the higher-than-expected scores from South African (especially) and Indian interviewees suggest the past British colonial history of those countries might represent a potential boundary condition impacting the effect of culture on British raters' evaluations. To further explore this, in Study 2, we again used interviewees from countries sharing a history with raters from the U.K. (i.e., Canada, South Africa, and India), but we replaced interviewees from Poland and Spain with interviewees from two alternative countries without a British colonial history: Chile and Germany.

We also explored the impact of interviewers' individual differences, such as SDO, RWA and ethnocentrism, on evaluations of interviewees from various countries. However, we failed to find any significant moderating effects of those constructs on the relationship between cultural differences and evaluations. While the within-subjects design of Study 1 ensured sufficient statistical power to examine the cultural differences, and our interaction effects were small and

far away from significance levels (e.g., η^2 ranging from .001 to .004), our sample size might have been too small to detect interactions with between-subject factors (see O'Boyle et al., 2019). As such, we increased our sample size in Study 2 to ensure larger statistical power to capture potential interactions.

Finally, Study 1 used the U.K. national average scores for cultural dimensions from the GLOBE research as the foundation for estimating cultural differences between raters and interviewees. Although this approach is common in cross-cultural research, it has also been criticized as an "ecological fallacy" (Brewer & Venaik, 2014). Indeed, it is possible that raters within the same country (and especially in a multicultural society like the U.K.) vary in their individual-level views and beliefs related to cultural dimensions (e.g., power distance or uncertainty avoidance). As such, Study 2 also included an individual-level measure of the GLOBE dimensions.

Study 2

Overall, Study 2 relied on a design similar as Study 1, with a few changes and additions to address key limitations of Study 1 and explore alternative explanations described above. First, we invited participants from the U.K. to evaluate interviewees from a slightly different set of five countries, replacing Poland and Spain with Germany and Chile, to further explore the shared colonial history explanation alongside cultural differences. Second, we selected a new set of ten interviewees (two per country), following an extensive pilot study. Third, we increased our sample size to help with testing potentially under-powered moderations. Finally, we removed our measure of SDO, which was highly correlated with RWA in Study 1, and replaced it with an individual-level measure of the GLOBE dimensions. Study 2 therefore allows us to further test

Hypotheses 1, 2, and 4, alongside alternative explanations. Table 5 shows the GLOBE's nine cultural dimensions value scores and composite scores, for the six countries relevant in Study 2.

Table 5 – Study 2 Country comparison of GLOBE cultural dimensions

	U.K.	Canada	Chile*	India	Germany	South Africa
Performance orientation	**5.90**	6.15	6.35	6.05	6.01	4.92
Assertiveness	**3.70**	4.15	3.25	4.76	3.09	3.82
Future orientation	**5.06**	5.35	5.78	5.60	4.85	5.20
Humane orientation	**5.43**	5.64	5.58	5.28	5.46	5.07
Institutional collectivism	**4.31**	4.17	5.32	4.71	4.82	4.30
In-group collectivism	**5.55**	5.97	6.15	5.32	5.18	4.99
Gender egalitarianism	**5.17**	5.11	4.98	4.51	4.89	4.26
Power distance	**2.80**	2.70	2.33	2.64	2.54	3.65
Uncertainty avoidance	**4.11**	3.75	4.66	4.73	3.32	4.79
Total distance score	**-**	**2.28**	**3.17**	**3.97**	**4.59**	**4.61**

Note. Based on GLOBE value scores. *Argentina value scores were used as proxy for Chile, since GLOBE has no data specific to Chile. Argentina is the country most similar to Chile politically, religiously, and ethnically (Tiano, 1986; Undurraga, 2015); as Chile was once connected to Argentina (Tulchin, 2010).

Methods

Sample

We recruited a total of 197 British participants with past hiring experience through Prolific. Our demographic variables in Study 2 were quite similar to those of Study 1. The mean age of participants was 43.5 (*SD* = 11.42). Participants were 45.5% female and 89.5% White (3% Black, 3.5% Asian, 4% Other). They had an average of 8.1 years of HR-related experience (*SD* = 9.48), 23.7% had experience using AVIs, and on average had lived 0.93 years abroad (i.e., outside of U.K. - *SD* = 2.82).

Procedure & Pilot Study

Participants followed the same procedure outlined in Study 1, with the following differences: (1) they watched recorded responses to three questions for each interviewee (vs. four in Study 1), (2) we used a different set of interviewees/videos from five countries selected from a

pilot study (see below), they completed an individual measure of culture created by the GLOBE project (House et al., 2004) instead of the SDO measure.

Prior to conducting Study 2, we pre-selected a new set of three video-responses from 18 interviewees from our five countries, who demonstrated equivalent objective performance (based on BARS scores from trained coders). We used the same process described in Study 1 for training coders and our ICCs for our two new countries (Chile, Germany) ranged from .74 to .87.

We created short clips with the first 30 seconds of their initial response, showed those videos to a sample of 50 U.S. residents recruited from Mechanical Turk, and asked them to assess each interviewee on measures of perceived warmth, competence, attractiveness, English proficiency, and accent. We ultimately selected two interviewees per country (one male, one female) who obtained similar scores on those elements, to be used in the main study (see our Online Supplement for more detail)[2].

Measures

We used the same AVI rating scale developed by Gorman et al. (2018) (α = .89 - .92), and individual measures of ethnocentrism (α = .88) and right-wing authoritarianism (α = .80) in Study 2. We also integrated the GLOBE (House et al., 2004) individual measure of cultural values, which includes 38 items measuring nine constructs. Items are rated on a 7-point scale. Example items include "In my country, people should be encouraged to be dominant (1) - non-dominant (7)" (assertiveness) or "I believe that power should be concentrated at the top (1) - shared throughout the organization (7)" (power distance). Please see the GLOBE project website for further details. Unfortunately, the internal consistency reliability for several cultural

---

[2] We also included examples of anonymized screenshots from the video responses in the online supplement.

dimensions ranged from acceptable to low: Humane Orientation (.79), Assertiveness (.69), Performance Orientation (.65), Future Orientation (.55), Power Distance (.53), Gender Egalitarianism (.53), In-group Collectivism (.46), Institutional Collectivism (.34), and Uncertainty Avoidance (.23). Because of the poor reliabilities for some dimensions, we do not report analyses using individual GLOBE measures as pure replacement to the national averages here. However, mean scores from the individual GLOBE measures were largely aligned with national averaged presented in Table 5 above. For instance, we found very similar scores for power distance (individual = 2.91 vs. national = 2.80) and uncertainty avoidance (3.81 vs. 4.11 – see online supplement for detailed individual scores for all dimensions).

Results

Preliminary analyses revealed one potential univariate outlier (Ethnocentrism, $z = 4.38$), but no severe item-response related issues. Removing the participant's data did not influence our results. Additionally, this participant passed our two attention checks embedded within our individual measures. Therefore, we decided to keep all 197 data points for our main analyses. Table 6 presents descriptive statistics and correlations among study variables for Study 2.

Table 6 – Correlations among study variables for Study 2

		M	SD	1	2	3	4	5	6	7	8
1	Age	43.49	11.45								
2	HR Experience	8.10	9.51	.34**							
3	RWA	2.49	0.71	.19**	.08						
4	Ethnocentrism	1.96	0.47	.17*	.02	.31**					
5	South Africa	3.37	0.82	-.11	-.03	.07	-.13				
6	India	3.02	0.77	-.20**	-.08	-.11	-.25**	.32**			
7	Canada	4.00	0.63	-.05	-.08	-.06	-.20**	.22**	.24**		
8	Germany	3.06	0.77	-.07	.01	-.04	.01	.31**	.34**	.24**	
9	Chile	3.16	0.68	-.01	-.09	-.15*	-.17*	.19**	.31**	.30**	.33**

Note. * $p < .05$, ** $p < .01$. India/South Africa/Canada/Germany/Chile variables refer to evaluation scores received by interviewee in those respective countries. HR Experience= years of experience in an HR role, RWA = right-wing authoritarianism

Cultural Differences

As noted in Table 5, we computed a total distance score for our two new countries (vs. the U.K.) from the GLOBE value scores, which resulted in the following predicting ranking, from most similar to dissimilar: Canada (2.28), Germany (3.17), India (3.97), Chile (4.59) and South Africa (4.61). To test Hypothesis 1, we again conducted a within-subjects repeated measure ANOVA to examine whether mean score differences across these five countries were statistically significant and aligned with the order predicted above. There was a statistically significant difference in mean scores across the five countries, $F(4, 788) = 81.781$, $p < .001$, $\eta^2 = .293$. Pairwise comparisons further showed that evaluations of interviewees from Canada ($M = 4.00$, $SD = 0.63$) were significantly higher than all other groups (all $p < .001$). Evaluations of interviewees from South Africa ($M = 3.37$, $SD = 0.82$) were significantly higher than those of Chile, Germany, and India (all $p < .001$). Chilean interviewees evaluations ($M = 3.16$, $SD = 0.68$) were significantly higher than evaluations of Indian ($M = 3.02$, $SD = 0.77$, $p = .023$), but not German ($M = 3.05$, $SD = 0.77$, $p = .078$) interviewees. And finally, German and Indian interviewees were also rated equivalently ($p = .601$). Overall, these findings, again, provide only partial support for Hypothesis 1.

Moderation Analyses

To test for the moderating role of ethnocentrism and RWA (Hypotheses 2 and 4), we standardized them and included them separately as covariates in repeated-measure ANCOVAs. Results are presented in Table 7. For ethnocentrism, in contrast to Study 1 but generally consistent with Hypothesis 2, the country of origin x ethnocentrism interaction was significant, $F(4, 784) = 2.569$, $p = .039$, $\eta^2 = .010$. Observing effects for raters high (vs. low) on ethnocentrism showed that differences in evaluations between countries were slightly larger for

more ethnocentric raters. This is also consistent with correlations in Table 6, showing negative relationships between ethnocentrism and ratings for interviewees across countries (except for Germany). In contrast, there was no significant country-of-origin x RWA interaction, thus not supporting Hypothesis 4 (and consistent with results from Study 1).

Table 7. Repeated-measure ANCOVA results for Study 2

	F	p	η^2	η_p^2
Country-of-origin	82.432	< .001	.293	.296
Ethnocentrism	10.276	.002	.010	.050
Country x ethnocentrism	2.569	.039	.009	.013
Country-of-origin	82.142	< .001	.293	.295
Right-Wing Authoritarianism	1.454	.229	.007	.007
Country x RWA	1.870	.116	.002	.009

Note. F-values are based on the Greenhouse-Geisser correction

Robustness checks

We controlled for participants' HR experience again to see whether biased evaluations against applicants from different cultures would be weaker (or stronger) for more experienced respondents. Our within-subjects repeated measure ANOVA using HR experience as a covariate did not reveal any significant interaction between country of origin and experience, $F(4, 784) = 0.459$, $p = .760$). But the main effect of the country of origin remained significant, $F(4, 784) = 48.590$, $p < .001$. We then explored whether our results could be caused by particularly low/high ratings obtained by one of our two applicants for each country using t-tests. We found no difference between Canadian $t(196) = 0.000$, $p = 1.00$, $d = 0.00$, Indian $t(196) = 0.336$, $p = .737$, $d = 0.05$, German $t(196) = 1.518$, $p = .131$, $d = 0.22$ interviewees. However, Chilean $t(196) = 2.136$, $p = .034$, $d = 0.30$, and South African $t(196) = 3.395$, $p = .001$, $d = 0.48$, interviewees were significantly different from one another. In both countries, the male interviewee received higher scores than the female interviewee.

Discussion

In Study 2, British participants specifically ranked Canadian interviewees highest, followed by South African, Chilean, German and lastly Indian interviewees. Despite controlling rigorously for interviewee performance as well as other potential confound variables via a pilot study (i.e., gender, perceived warmth, competence, attractiveness, English proficiency), a clear preference for Canadian (and to a lesser extent South African) interviewees over interviewees from other countries was observed. We therefore only found partial support for our prediction based on the GLOBE framework of how cultural distance should impact evaluations. Particularly the South African interviewees received noticeably higher-than-expected evaluations. These findings are also not fully consistent with prior work based on the similarity-attraction theory (i.e., Montoya & Horton, 2013). Indeed, according to GLOBE values, the British culture is less similar to the South African than the Chilean or Germanic cultures, but South African interviewees were rated more positively than Chilean or Germanic ones. In addition, all our South African interviewees were Black (vs. Germans who were White, like the vast majority of our raters), suggesting that race did not play a central role (at least not a negative one). Although the higher-than-expected ratings for South African interviewees are consistent with those found in Study 1, we did not find the same effect for Indian interviewees, therefore largely eliminating the shared colonial history suggested in Study 1 as a viable explanation for our findings.

We also note that gender bias may have played a role in how evaluations in this study were assigned. Although evaluations assigned to male and female interviewees within Canada, India and Germany were not statistically different, male interviewees were rated significantly higher than female interviewees within South Africa and Chile. This finding aligns with the work based on the stereotype content model (e.g., Fiske et al., 2002) that often shows female

professionals are perceived as less competent that males (despite similar qualifications - see Cuddy et al., 2008).

We also re-examined the potential role of ethnocentrism and RWA as two individual differences that could moderate how evaluations of interviewees from culturally-distant countries were assigned. Given that moderation effects can be more difficult to detect (O'Boyle et al., 2019), we relied on a larger sample to increase statistical power. Like in Study 1, RWA did not play a moderating role, despite past research showing that people high on RWA are more likely to favor in-group versus out-group members (Altemeyer, 1988). However, in line with our predictions but contrary to Study 1, we found that ethnocentrism played a (small) moderating role, with differences in ratings assigned to more culturally-distant interviewees being more pronounced for raters higher on ethnocentrism. This could be because high-RWA individuals are concerned about protecting their in-group against any kind of out-group-based threat, but less so about different categories of out-group members. In contrast, ethnocentrism is more directly connected to one's level of openness to cultures that differ from their own (Neuliep & McCroskey, 1997).

Unfortunately, our individual measure of culture (extracted from the GLOBE project webpage) had poor reliability coefficients for most constructs. Humane Orientation, Performance Orientation and Assertiveness had acceptable reliabilities whereas the remaining six construct reliabilities were weak. This prevented us from conducting more interesting analyses, for instance by looking at how the cultural distance between raters' individual GLOBE scores and the GLOBE national scores for our five countries of interest could impact evaluations of interviewees. However, this data helped us confirm that individual-level mean scores national-level scores from the GLOBE project were largely similar.

General Discussion

The present research investigated how cultural difference based on country of origin could bias interviewers' evaluations in an AVI context. We proposed that evaluators from one culture should evaluate more positively applicants from cultures closer to theirs and more negatively applicants from cultures more distant to theirs. We tested these predictions in two separate studies with similar designs in the U.K. context, examining how British participants with hiring experience assigned evaluations to interviewees from five culturally distinct countries. We relied on the GLOBE framework (House et al., 2004) and similarity-attraction theory (i.e., Montoya & Horton, 2013) to predict how evaluation rankings would be assigned. We found significant differences in how evaluations were assigned to interviewees, thus offering initial evidence for the presence of potential cultural biases in an AVI context. The ranking order of evaluation scores assigned to interviewees were mostly consistent with our GLOBE framework predictions, with a couple of noticeable exceptions. For example, using the GLOBE framework we predicted that British participants would perceive Canadian interviewees as being most 'similar' and South African interviewees as most 'dissimilar'. Although Canadian interviewees did consistently receive high evaluations, surprisingly so did our South African interviewees. Further, the evaluations of interviewees from other countries (e.g., Spain and Poland in Study 1, Chile and Germany in Study 2) were generally consistent with the GLOBE-based predictions, but ratings of Indian participants were higher than expected in Study 1.

Overall, our findings suggest that, while evaluators' judgments of interviewees were indeed influenced by their cultural similarity, other contextual factors may have also played a role. For example, within the context of current global affairs, the surprisingly-high ratings obtained by South African interviewees could have been influenced by social justice movements

like 'Black Lives Matter'. All South African interviewees in both studies were Black. Thus, such movements may have increased participants' sensitivity or awareness of the issue of bias, especially towards Black individuals, explaining why South African interviewees received among the highest ratings. This positive effect could be specific to the AVI context, or represent indirect preliminary evidence that social justice movements are helping change views of minority groups (at least of Black individuals for U.K. raters).

A possible explanation for the higher evaluations received by Indians participants in Study 1 vs. Study 2 could relate to India's unique caste system, and how it can be associated with (perceived) higher vs. lower skills in migrant populations. The Indian community makes up the largest minority group in the U.K. (with some 1.6 million members) and have been interwoven into British society for several generations. However, while certainly not as prevalent as in India, and not recognized as a legally-protected group (Waughray, 2009), a caste system does also exist in the U.K.. For example, up to 200,000 Indian immigrants in the U.K. are considered *dalits*, formerly known as 'untouchables' (Blunt, 2010; Modood et al., 1997). The literature on discrimination highlights how attitudes towards migrant groups traditionally viewed as low-skilled are generally more negative than towards migrant groups traditionally viewed as highly skilled (Hainmueller et al., 2015). While carefully examining the specific ratings for the different Indian interviewees used in our two studies, we noted that interviewees in Study 1 were evaluated generally favorably on constructs relevant to highly-skilled roles: perceived intelligence, mental skills (with means ranging from 3.6 to 3.8 out of 5) as well as leadership (3.2). In contrast, evaluations were somewhat lower for those constructs for Indian interviewees in Study 2, with means of 3.0-3.1 for intelligence and mental skills or 2.6 for leadership. It is

thus possible that our U.K. participants *perceived* the Indian interviewees in Study 1 as highly skilled migrants but those in Study 2 as less-skilled migrants.

In addition, our data collection was conducted shortly after Brexit, a major political shift in the U.K.'s position within the E.U. that, among many other things, captured British voters' sentiment towards immigrants and immigration. In fact, specific accounts exist in the literature highlighting British racism and negative stereotypes or attitudes towards Polish (Rzepnikowska, 2019), as well as German migrants (Grix & Lacroix, 2006; Wittlinger, 2004). This might explain why evaluation of the Polish applicants in Study 1 were slightly more negative than predicted by the GLOBE cultural scores. And finally, a recent field experiment conducted by Heath and Di Stasio (2019) examining invitations for a job interview for applicants of different national origins revealed that South American (i.e., Latino) immigrants were among the least likely to be invited for interviews compared to other immigrant groups in the U.K. Despite portraying characteristics of warmth, competence, and possessing strong English skills, British interviewees may qualify this ethnic group as being less suitable for employment. This could help explain the consistently lower ratings (although mostly aligned with our GLOBE predictions) that Spanish and Chilean interviewees received across both our studies.

Beyond purely culture-based explanations, we also considered the role that accents can play in impacting interviewee evaluations. The accent literature suggests that interviewees whose accents are perceived as 'foreign' or non-native (Munro & Derwing, 1995) should receive lower evaluations (Campbell-Kibler, 2007; Lippi-Green, 2012). However, and surprisingly, accents did not seem to significantly impact how evaluations were assigned to interviewees in both our studies. For example, the Indian interviewees in Study 1 received amongst the highest evaluations across all five countries, despite having arguably the strongest accents (certainly

stronger than the Canadian, Polish, and South African interviewees). In Study 2, the South African and Chilean interviewees received higher evaluations than German interviewees despite having accents that were perceived to be slightly stronger (in our pilot study). Additionally, accent-biased evaluations can also depend on accent 'type', for instance French accent being viewed positively but Spanish accent negatively (see Deprez-Sims & Morris, 2010; Hosoda et al., 2012). However, our results suggest that other factors (like culture) are likely at play, since Spaniards received poor evaluations in Study 1, whereas Chilean interviewees (whose first language is also Spanish, and arguably have a similar Latino accent) received relatively average evaluations in Study 2. Overall, we believe our efforts to carefully select interviewees with similar levels of English proficiency likely reduced the effect of accents on evaluations across both our studies – as there was no clear observable pattern to support that accents played a significant role in the manner that evaluations were assigned. In other words, the interviewees with 'foreign' perceived accents likely did not fall victim to accent-related discrimination because they were still able to communicate 'effectively' (Creese & Kambere, 2003; Hideg et al., 2022).

Our findings also differ from previous research on (American) racial hierarchy. For instance, Bell et al. (2014) showed that White (U.S.) interviewers systematically prefer not only White over African-American applicants but also White immigrants (e.g., from Western Europe) over Black immigrants (e.g., from Africa). Although the current research was designed to specifically examine the role of culture (instead of race), our findings suggest that in the context of international/cross-cultural hiring, culture could be more relevant than race. Indeed, across both our studies South African interviewees (i.e., the only Black interviewees) received amongst the highest ratings. In contrast, Polish (Study 1) and German (Study 2) interviewees, despite

being White, received significantly lower evaluations. In addition, when discussing our Study 1

results, we initially argued that the shared colonial history between the U.K. and Canada, South

Africa, and India could explain our pattern of findings (and especially the higher-than-expected

ratings for the latter two countries). However, the findings of Study 2 were not fully consistent

with that explanation, given the lower ratings observed for Indian interviews.

Finally, we also explored the potential moderating role of individual-level measures of

prejudice across both studies. Unlike previous research (Sibley & Duckitt, 2008), we did not find

RWA and SDO to be powerful predictors of prejudice in our two studies. However, levels of

ethnocentrism did significantly (but modestly) influence how British participants evaluated

interviewees from other countries in Study 2. It could also be that the weak interactions observed

for RWA or SDO could be due to generally lower scores obtained in our samples, perhaps

because of social desirability (i.e., with 'seasoned' Prolific participants used to research surveys

being able to detect the constructs being measured and respond accordingly). That said, our

findings highlight potential boundary conditions to the role played by such individual

differences. For example, while SDO or RWA have been shown to impact how people from one

country view outgroup members (e.g., discriminates against immigrants or ethnic minority

groups – see Sibley & Duckitt, 2008 or Sidanius & Pratto, 1999), our findings suggest that such

individual difference might work differently when one is comparing and judging individuals

from multiple out-groups, at least in the context of AVIs. In other words, RWA and SDO might

be more relevant when examining simpler in-group vs. out-group judgments, whereas

ethnocentrism could become somewhat more important when comparing multiple (culturally

diverse) out-groups.

Practical Implications

This research has important practical implications for cross-cultural hiring, particularly in an AVI context, which is highly relevant as organizations increasingly rely on virtual interviewing tools. Hiring managers and HR professionals have significant influence over the design and outcome of the selection process. Our most significant implication for practitioners is our finding(s) across both studies that AVIs may be subject to discrimination/bias based on country of origin. For example, our research suggests there could be biases against applicants from different cultures, even in a structured AVI context - and especially if raters are relying on subjecting rating systems (i.e., like the Gorman AVI construct measure we used in our studies). Despite controlling rigorously for interviewee performance, both of our studies found significant differences in how mean evaluation scores were assigned to interviewees from various countries. In other words, HR raters using AVIs to recruit internationally under corporately mandated equity, diversity, and inclusion (EDI) objectives to increase their workforce diversity, may fail to achieve to do so if monocultural selection committees are used. To counteract this potential pitfall, we recommend practitioners to carefully select culturally-heterogeneous selection committees (when possible) to avoid evaluations influenced by 'similar to me biases' (Kristof-Brown et al., 2002) and similarity-attraction effects (Montoya & Horton, 2013).

Our second practical implication is that our studies indirectly demonstrated the importance of using standardized ratings like BARS and training raters to be unbiased in selection processes. For example, we were able to train several coders on how to use a BARS to objectively assign evaluations to diverse interviewees. We specifically used standardized rating scales vs. the unstructured rating measures used by our raters which more closely reflects how these processes are used frequently in practice (i.e., where raters are often asked to rate simply on a 1-5 stars system or the like). As such, our findings provide additional evidence for the value

of using standardized rating scales like BARS (i.e., one of the key elements of structured interviews – see Levashina et al., 2014 for a review) as part of the evaluation process in AVIs.

Limitations and Future Research Directions

This research has a number of limitations and associated suggestions for future research. First, while our approach to examine the impact of cultural differences is similar to past research (e.g., using the GLOBE framework), it is possible that there are also individual-level differences. In other words, not all British interviewers are exactly similar in terms of their cultural values, and also may not be fully consistent with GLOBE's aggregated findings for British values. Such an approach has been critiqued in the literature (e.g., Brewer & Venaik, 2014). While we attempted to include an individual-level GLOBE measure in Study 2, the poor reliabilities for many dimensions made it impossible to use that data for conducting alternative analyses. This might signal the need to develop a more psychometrically-sound measure of the nine GLOBE dimensions. We also acknowledge the limitation of using overall cultural distance scores computed by the sum of the nine GLOBE cultural dimensions to formulate and test our first hypothesis. While we attempted to derive alternative predictions by assigning theoretically-derived weights to different cultural dimensions (vs. unit weighting used originally), the revised predicted country ranking did not differ substantially from those associated to Hypothesis 1, and did not align more closely with our findings (see online supplement for more details). Future research should consider alternative ways of varying 'weightings' that could be assigned to cultural dimensions that are deemed as more/less relevant to an interview context (i.e., performance orientation).

Second, both our studies relied on a within-subjects design where participants watched only one interviewee per country (and not all possible interviewees), thus leading to "missing-

by-design" data limiting analytical options. This approach was chosen because it closely

replicates how hiring decisions are made in organizational settings (i.e., where raters evaluate

multiple applicants – possibly with different cultural backgrounds - for one position). In addition,

we carefully selected videos of real interviewees from a separate study, which were rated

similarly on behaviourally-anchored rating scales (BARS), and took additional steps to ensure

that candidates were also equivalent on other elements (e.g., warmth, competence, attractiveness,

accents). This approach was used to increase the external validity of our study. Yet, some of our

selected interviewees received significantly higher evaluations than their respective country

counterparts (e.g., male vs. female South African interviewees in Study 2). To avoid such issues,

future studies could use actors reading scripts, which would increase internal validity (but

decrease external validity and generalizability).

Third, we used an experimental design with online panel participants from one country

(the U.K.) assessing videos interviewees from a limited number of countries completing mock

interviews. Future research should replicate this study using real job interviews, where stakes are

higher both for interviewers and applicants, which might bolster the effect of cultural

differences. Future studies could also examine evaluators from different country of origin to see

if similar or different findings emerge. Such studies could be used to further explore the shared

cultural history argument presented in our Study 1 discussion (i.e., by selecting raters from a

country with no colonial history with the candidates' countries of origin).

Fourth, we focused on AVIs, as a novel and unique form of interviews. Yet, it remains

unclear what kind of biases exists in AVIs (e.g., Lukacik et al., 2020) and exactly how they differ

from those present in-person interviews. The AVI format might partly explain the non-

significant findings for our moderators. Indeed, RWA, SDO, or ethnocentrism might be activated

specifically when the in-group vs. out-group nature such as in interpersonal interactions. As such, their effects on evaluations could have been stronger in face-to-face interviews. Alternatively, future research could also explore other individual difference measures suitable for an interview context such as the Cultural Tolerance scale (Gasser & Tan, 1999) or Intercultural Willingness to Communicate scale (Kassing, 1997).

Finally, our study relied on one job (a management associate in an international bank). But future studies could examine whether the role of culture, but also race or accents, varies depending on job type, job level, or industry. For instance, Hosoda et al. (2012) found applicant with a Mexican-Spanish accent received lower evaluations for entry-level engineering job. However, Timming (2017) showed that the position for which non-native accent applicants apply for (i.e., customer facing versus non-customer facing) could play a role. The position used in our studies included a strong customer service orientation, which might have similarly enhanced the effect of culture on ratings. As such, future studies could explore if the effect of cultural distance between interviewees and raters in AVIs also depend on job type. Alternatively, future research could manipulate the industry, job type, or level to see if some accents are more favourably evaluated than others in the context of an AVI.

Conclusion

This research represents a first attempt to empirically examine potential cultural biases in an AVI context. We specifically explored how British interviewers evaluated interviewees from a total of seven different cultures. We also examined three individual measures of prejudice that could potentially moderate these evaluations (i.e., ethnocentrism, RWA, and SDO). Importantly, evaluations of interviewees were only partially consistent with GLOBE-based cultural distance between their country of origin and the interviewers'. Instead, we found that interviewees

Canada and South Africa received consistently higher evaluations than did those interviewees from other countries across both our studies. These findings have particular significance for hiring professionals engaged in online interviewing, where (unintentional) biases related to applicant country of origin may impact how evaluations are assigned. With the recent increased usage of virtual tools (due to the COVID pandemic), organizations need to be aware of how changes in their hiring processes impact selection decisions.

References

Allen, D. G., & Vardaman, J. M. (2017). Recruitment and retention across cultures. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*(1), 153–181. https://doi.org/10.1146/annurev-orgpsych-032516-113100

Altemeyer, B. (1988). *Enemies of Freedom: Understanding right-wing authoritarianism.* Jossey-Bass.

Arseneault, R., & Roulin, N. (2021). A theoretical model of cross-cultural impression management in employment interviews. *International Journal of Selection and Assessment*, *29*(3–4), 352–366. https://doi.org/10.1111/ijsa.12348

Bartkoski, T., Herrmann, E., Witt, C., & Rudolph, C. W. (2018). *A meta-analysis of employment discrimination against muslims and arabs*. PsyArXiv. https://doi.org/10.31234/osf.io/6nw9h

Basch, J. M., & Melchers, K. G. (2021). The use of technology-mediated interviews and their perception from the organization's point of view. *International Journal of Selection and Assessment*, *n/a*(n/a). https://doi.org/10.1111/ijsa.12339

Basch, J. M., Melchers, K. G., Kegelmann, J., & Lieb, L. (2020). Smile for the camera! The role of social presence and impression management in perceptions of technology-mediated

interviews. *Journal of Managerial Psychology*, *35*(4), 285–299.

https://doi.org/10.1108/JMP-09-2018-0398

Basch, J. M., Melchers, K. G., Kurz, A., Krieger, M., & Miller, L. (2021). It takes more than a

good camera: Which factors contribute to differences between face-to-face interviews

and videoconference interviews regarding performance ratings and interviewee

perceptions? *Journal of Business and Psychology*, *36*(5), 921–940.

https://doi.org/10.1007/s10869-020-09714-3

Bell, M.P., Marquardt, D., & P. Berry, D. (2014). "Diversity," immigration, and the new

American multi-racial hierarchy. *Journal of Managerial Psychology*, *29*(3), 285–303.

https://doi.org/10.1108/JMP-08-2012-0242

Bizumic, B., & Duckitt, J. (2018). Investigating right wing authoritarianism with a very short

authoritarianism scale. *Journal of Social and Political Psychology*, *6*(1), Article 1.

https://doi.org/10.5964/jspp.v6i1.835

Blacksmith, N., Willford, J., & Behrend, T. (2016). Technology in the employment interview: A

meta-analysis and future research agenda. *Personnel Assessments and Decisions*, *2*, 12–

20. https://doi.org/10.25035/pad.2016.002

Blunt, S. E. (2010). *The Caste System of Northern India*. Gyan Publishing House.

Brenner, F. S., Ortner, T. M., & Fay, D. (2016). Asynchronous video interviewing as a new

technology in personnel selection: The applicant's point of view. *Frontiers in

Psychology*, *7*, 863. https://doi.org/10.3389/fpsyg.2016.00863

Brewer, P., & Venaik, S. (2011). Individualism–Collectivism in Hofstede and GLOBE. *Journal

of International Business Studies*, *42*(3), 436–445.

Brewer, P., & Venaik, S. (2014). The ecological fallacy in national culture research. *Organization Studies*, *35*(7), 1063–1086. https://doi.org/10.1177/0170840613517602

Brunt, C. (2016). *Human Resource Management in International NGOs*. Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-57306-3

Cable, D. M., & Judge, T. A. (1997). Interviewers' perceptions of person–organization fit and organizational selection decisions. *Journal of Applied Psychology*, *82*(4), 546–561. https://doi.org/10.1037/0021-9010.82.4.546

Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, *82*(1), 32–64. https://doi.org/10.1215/00031283-2007-002

Cargile, A. C. (2000). Evaluations of employment suitability: Does accent always matter? *Journal of Employment Counseling*, *37*(3), 165–177. https://doi.org/10.1002/j.2161-1920.2000.tb00483.x

Carnevale, J. B., & Hatak, I. (2020). Employee adjustment and well-being in the era of COVID-19: Implications for human resource management. *Journal of Business Research*, *116*, 183–187. https://doi.org/10.1016/j.jbusres.2020.05.037

Chamorro-Premuzic, T., Winsborough, D., Sherman, R. A., & Hogan, R. (2016). New talent signals: Shiny new objects or a brave new world? *Industrial and Organizational Psychology*, *9*(3), 621–640. https://doi.org/10.1017/iop.2016.6

Charles-Toussaint, G. C., & Crowson, H. M. (2010). Prejudice against international students: The role of threat perceptions and authoritarian dispositions in US students. *The Journal of Psychology*, *144*(5), 413–428.

Creese, G., & Kambere, E. N. (2003). "What colour is your English?"*. *Canadian Review of Sociology/Revue canadienne de sociologie*, *40*(5), 565–573. https://doi.org/10.1111/j.1755-618X.2003.tb00005.x

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 61–149). Academic Press. https://doi.org/10.1016/S0065-2601(07)00002-0

Deprez-Sims, A.-S., & Morris, S. B. (2010). Accents in the workplace: Their effects during a job interview. *International Journal of Psychology*, *45*(6), 417–426. https://doi.org/10.1080/00207594.2010.499950

Deprez-Sims, A.-S., & Morris, S. B. (2013). The effect of non-native accents on the evaluation of applicants during an employment interview: The development of a path model. *International Journal of Selection and Assessment*, *21*(4), 355–367. https://doi.org/10.1111/ijsa.12045

Derous, E. (2017). Ethnic minorities' impression management in the interview: Helping or hindering? *Frontiers in Psychology*, *8*, 86. https://doi.org/10.3389/fpsyg.2017.00086

Derous, E., Buijsrogge, A., Roulin, N., & Duyck, W. (2016). Why your stigma isn't hired: A dual-process framework of interview bias. *Human Resource Management Review*, *26*(2), 90–111. https://doi.org/10.1016/j.hrmr.2015.09.006

Derous, E., Ryan, A. M., & Nguyen, H.-H. D. (2012). Multiple categorization in resume screening: Examining effects on hiring discrimination against Arab applicants in field and lab settings. *Journal of Organizational Behavior*, *33*(4), 544–570. https://doi.org/10.1002/job.769

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science*, *11*(4), 315–319. https://doi.org/10.1111/1467-9280.00262

Duckitt, J., & Sibley, C. (2017). *The dual process motivational model of ideology and prejudice* (pp. 188–221).

Dunlop, P. D., Holtrop, D., & Wee, S. (2022). How asynchronous video interviews are used in practice: A study of an Australian-based AVI vendor. *International Journal of Selection and Assessment*. https://doi.org/10.1111/ijsa.12372

Farndale, E., & Sanders, K. (2017). Conceptualizing HRM system strength through a cross-cultural lens. *The International Journal of Human Resource Management*, *28*(1), 132–148. https://doi.org/10.1080/09585192.2016.1239124

Fell, C. B., & König, C. J. (2016). Cross-cultural differences in applicant faking on personality tests: A 43-nation study. *Applied Psychology*, *65*(4), 671–717. https://doi.org/10.1111/apps.12078

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Gasser, M. B., & Tan, R. N. (1999). Cultural tolerance: Measurement and latent structure of attitudes toward the cultural practices of others. *Educational and Psychological Measurement*, *59*(1), 111–126. https://doi.org/10.1177/0013164499591008

Gorman, C. A., Robinson, J., & Gamble, J. S. (2018). An investigation into the validity of asynchronous web-based video employment-interview ratings. *Consulting Psychology Journal: Practice and Research*, *70*(2), 129. https://doi.org/10.1037/cpb0000102

Griswold, K. R., Phillips, J. M., Kim, M. S., Mondragon, N., Liff, J., & Gully, S. M. (2021). Global differences in applicant reactions to virtual interview synchronicity. *The International Journal of Human Resource Management*, *0*(0), 1–28. https://doi.org/10.1080/09585192.2021.1917641

Grix, J., & Lacroix, C. (2006). Constructing Germany's image in the British press: An empirical analysis of stereotypical reporting on Germany. *Journal of Contemporary European Studies*, *14*(3), 373–392. https://doi.org/10.1080/14782800601102617

Guchait, P., Ruetzler, T., Taylor, J., & Toldi, N. (2014). Video interviewing: A potential selection tool for hospitality managers – A study to understand applicant perspective. *International Journal of Hospitality Management*, *36*, 90–100. https://doi.org/10.1016/j.ijhm.2013.08.004

Hainmueller, J., Hiscox, M. J., & Margalit, Y. (2015). Do concerns about labor market competition shape attitudes toward immigration? New evidence. *Journal of International Economics*, *97*(1), 193–207. https://doi.org/10.1016/j.jinteco.2014.12.010

Hansen, K., & Dovidio, J. F. (2016). Social dominance orientation, nonnative accents, and hiring recommendations. *Cultural Diversity & Ethnic Minority Psychology*, *22*(4), 544–551. https://doi.org/10.1037/cdp0000101

Heath, A. F., & Di Stasio, V. (2019). Racial discrimination in Britain, 1969–2017: A meta-analysis of field experiments on racial discrimination in the British labour market. *The British Journal of Sociology*, *70*(5), 1774–1798. https://doi.org/10.1111/1468-4446.12676

Hideg, I., Shen, W., & Hancock, S. (2022). What is that I hear? An interdisciplinary review and research agenda for non-native accents in the workplace. *Journal of Organizational Behavior*, *43*(2), 214–235. https://doi.org/10.1002/job.2591

Hiemstra, A. M. F., Oostrom, J. K., Derous, E., Serlie, A. W., & Born, M. Ph. (2019). Applicant perceptions of initial job candidate screening with asynchronous job interviews: Does personality matter? *Journal of Personnel Psychology*, *18*(3), 138–147. https://doi.org/10.1027/1866-5888/a000230

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., Foels, R., & Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO_7 scale. *Journal of Personality and Social Psychology*, *109*(6), 1003–1028. https://doi.org/10.1037/pspi0000033

Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, *10*(4), 15–41.

Hosoda, M., Nguyen, L. T., & Stone-Romero, E. F. (2012). The effect of Hispanic accents on employment decisions. *Journal of Managerial Psychology*, *27*(4), 347–364. https://doi.org/10.1108/02683941211220162

House, R., Hanges, P., Javidan, M., Dorfman, P., & Gupta, V. (2004). *Leadership and Organizations: The GLOBE Study of 62 Societies*.

Huffcutt, A. I., & Arthur, W. (1994). *Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs.*

Huffcutt, A. I., Van Iddekinge, C. H., & Roth, P. L. (2011). Understanding applicant behavior in employment interviews: A theoretical model of interviewee performance. *Human Resource Management Review*, *21*(4), 353–367. https://doi.org/10.1016/j.hrmr.2011.05.003

Javidan, M., Dorfman, P., Sully de Luque, M., & House, R. (2006). In the eye of the beholder: Cross cultural lessons in leadership from project GLOBE. *Academy of Management Perspectives, The*, *20*, 67–90. https://doi.org/10.5465/AMP.2006.19873410

Kassing, J. W. (1997). Development of the intercultural willingness to communicate scale. *Communication Research Reports*, *14*(4), 399–407. https://doi.org/10.1080/08824099709388683

Kith, S., Cassidy, S. A., & Powell, D. M. (2022). The structured interview's resistance to gender discrimination under cognitive load. *International Journal of Selection and Assessment*, *30*(2), 281–301. https://doi.org/10.1111/ijsa.12361

Krings, F., Sczesny, S., & Kluge, A. (2011). Stereotypical inferences as mediators of age discrimination: The role of competence and warmth. *British Journal of Management*, *22*. https://doi.org/10.1111/j.1467-8551.2010.00721.x

Kristof-Brown, Barrick, M. R., & Franke, M. (2002). Applicant impression management: Dispositional influences and consequences for recruiter perceptions of fit and similarity. *Journal of Management*, *28*(1), 27–46. https://doi.org/10.1016/S0149-2063(01)00131-3

Kristof-Brown, & Jansen, K. J. (2007). *Issues of person-organization fit.*

Langer, M., König, C. J., & Fitili, A. (2018). Information as a double-edged sword: The role of computer experience and information on applicant reactions towards novel technologies for personnel selection. *Computers in Human Behavior*, *81*, 19–30. https://doi.org/10.1016/j.chb.2017.11.036

Langer, M., König, C., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment*, *25*. https://doi.org/10.1111/ijsa.12191

Lee, I.-C., Pratto, F., & Johnson, B. T. (2011). Intergroup consensus/disagreement in support of group-based hierarchy: An examination of socio-structural and psycho-cultural factors. *Psychological Bulletin*, *137*(6), 1029–1064. https://doi.org/10.1037/a0025410

Levashina, J., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology*, *67*(1), 241–293. https://doi.org/10.1111/peps.12052

Lin, Y., Rancer, A. S., & Trimbitas, O. (2005). Ethnocentrism and intercultural-willingness-to-communicate: A cross-cultural comparison between Romanian and US American college students. *Journal of Intercultural Communication*, *34*(2), 138–151.

Lippi-Green, R. (2012). *English with an accent: Language, ideology, and discrimination in the United States.* Routledge.

Lukacik, E.-R., Bourdage, J. S., & Roulin, N. (2020). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review*, 100789. https://doi.org/10.1016/j.hrmr.2020.100789

Lukacik, E.-R., Bourdage, J. S., & Roulin, N. (2022). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review*, *32*(1), 100789. https://doi.org/10.1016/j.hrmr.2020.100789

Manroop, L., Boekhorst, J. A., & Harrison, J. A. (2013). The influence of cross-cultural differences on job interview selection decisions. *The International Journal of Human Resource Management*, *24*(18), 3512–3533. https://doi.org/10.1080/09585192.2013.777675

Maurer, R., & Maurer, R. (2020). Job interviews go virtual in response to COVID-19. *Society for Human Resources Management*. https://www.shrm.org/resourcesandtools/hr-topics/talent-acquisition/pages/job-interviews-go-virtual-response-covid-19-coronavirus.aspx

McCarthy, J. M., Iddekinge, C. H. V., & Campion, M. A. (2010). Are highly structured job interviews resistant to demographic similarity effects? *Personnel Psychology*, *63*(2), 325–359. https://doi.org/10.1111/j.1744-6570.2010.01172.x

McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith - a failure of analysis. *Human Relations*, *55*(1), 89–118. https://doi.org/10.1177/0018726702551004

Mejia, C., & Torres, E. N. (2018). Implementation and normalization process of asynchronous video interviewing practices in the hospitality industry. *International Journal of Contemporary Hospitality Management*, *30*(2), 685–701. https://doi.org/10.1108/IJCHM-07-2016-0402

Modood, T., Berthoud, R., Lakey, J., Nazroo, J., Smith, P., Virdee, S., & Beishon, S. (1997). *Ethnic minorities in Britain: Diversity and disadvantage*. Policy Studies Institute.

Montoya, R. M., & Horton, R. S. (2013). A meta-analytic investigation of the processes underlying the similarity-attraction effect. *Journal of Social and Personal Relationships*, *30*(1), 64–94. https://doi.org/10.1177/0265407512452989

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*(3), 289–306. https://doi.org/10.1177/002383099503800305

Nesdale, A. R., & Rooney, R. (1990). Effect of children's ethnic accents on adults' evaluations and stereotyping. *Australian Journal of Psychology*, *42*(3), 309–319. https://doi.org/10.1080/00049539008260128

Neuliep, J. W., Chaudoir, M., & McCroskey, J. C. (2001). A cross-cultural comparison of ethnocentrism among Japanese and United States college students. *Communication Research Reports*, *18*(2), 137–146. https://doi.org/10.1080/08824090109384791

Neuliep, J. W., Hintz, S. M., & McCroskey, J. C. (2005). The influence of ethnocentrism in organizational contexts: Perceptions of interviewee and managerial attractiveness, credibility, and effectiveness. *Communication Quarterly*, *53*(1), 41–56. https://doi.org/10.1080/01463370500055954

Neuliep, J. W., & McCroskey, J. C. (1997). The development of a U.S. and generalized ethnocentrism scale. *Communication Research Reports*, *14*(4), 385–398. https://doi.org/10.1080/08824099709388682

Neuliep, J. W., & Speten-Hansen, K. M. (2013). The influence of ethnocentrism on social perceptions of nonnative accents. *Language & Communication*, *33*(3), 167–176. https://doi.org/10.1016/j.langcom.2013.05.001

O'Boyle, E., Banks, G. C., Carter, K., Walter, S., & Yuan, Z. (2019). A 20-year review of outcome reporting bias in moderated multiple regression. *Journal of Business and Psychology*, *34*(1), 19–37. https://doi.org/10.1007/s10869-018-9539-8

Petersen, L.-E., & Dietz, J. (2000). Social discrimination in a personnel selection context: The effects of an authority's instruction to discriminate and followers' authoritarianism.

Journal of Applied Social Psychology, *30*(1), 206–220. https://doi.org/10.1111/j.1559-1816.2000.tb02312.x

Petersen, L.-E., & Dietz, J. (2005). Prejudice and enforcement of workforce homogeneity as explanations for employment discrimination. *Journal of Applied Social Psychology*, *35*(1), 144–159. https://doi.org/10.1111/j.1559-1816.2005.tb02097.x

Pogrebtsova, E., Luta, D., & Hausdorf, P. A. (2020). Selection of gender-incongruent applicants: No gender bias with structured interviews. *International Journal of Selection and Assessment*, *28*(1), 117–121. https://doi.org/10.1111/ijsa.12270

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, *114*(41), 10870–10875. https://doi.org/10.1073/pnas.1706255114

Roulin, N., Lukacik, E.-R., Bourdage, J. S., Clow, L., Bakour, H., & Diaz, P. (2022). Bias in the background? The role of background information in asynchronous video interviews. *Journal of Organizational Behavior*, *n/a*(n/a). https://doi.org/10.1002/job.2680

Rzepnikowska, A. (2019). Racism and xenophobia experienced by Polish migrants in the UK before and after Brexit vote. *Journal of Ethnic and Migration Studies*, *45*(1), 61–77. https://doi.org/10.1080/1369183X.2018.1451308

Sandal, G. M., van de Vijver, F., Bye, H. H., Sam, D. L., Amponsah, B., Cakar, N., Franke, G. H., Ismail, R., Kjellsen, K., Kosic, A., Leontieva, A., Mortazavi, S., & Tien-Lun Sun, C. (2014). Intended self-presentation tactics in job interviews: A 10-country study. *Journal*

of Cross-Cultural Psychology, *45*(6), 939–958.

https://doi.org/10.1177/0022022114532353

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel

psychology: Practical and theoretical implications of 85 years of research findings.

Psychological Bulletin, *124*(2), 262–274. https://doi.org/10.1037/0033-2909.124.2.262

Schwartz, S. (2006). A theory of cultural value orientations: Explication and applications.

Comparative Sociology, *5*(2–3), 137–182.

Scullion, H., Vaiman, V., & Collings, D. G. (2016). Strategic talent management. *Employee*

Relations, *38*(1), 1–7.

Sears, G. J., & Rowe, P. M. (2003). A personality-based similar-to-me effect in the employment

interview: Conscientiousness, affect-versus competence-mediated interpretations, and the

role of job relevance. *Canadian Journal of Behavioural Science / Revue Canadienne Des*

Sciences Du Comportement, *35*(1), 13–24. https://doi.org/10.1037/h0087182

Segrest Purkiss, S. L., Perrewé, P. L., Gillespie, T. L., Mayes, B. T., & Ferris, G. R. (2006).

Implicit sources of bias in employment interview judgments and decisions.

Organizational Behavior and Human Decision Processes, *101*(2), 152–167.

https://doi.org/10.1016/j.obhdp.2006.06.005

Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical

review. *Personality and Social Psychology Review: An Official Journal of the Society for*

Personality and Social Psychology, Inc, *12*(3), 248–279.

https://doi.org/10.1177/1088868308319226

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression* (pp. x, 403). Cambridge University Press. https://doi.org/10.1017/CBO9781139175043

Stone, D. L., Deadrick, D. L., Lukaszewski, K. M., & Johnson, R. (2015). The influence of technology on the future of human resource management. *Human Resource Management Review*, *25*(2), 216–231. https://doi.org/10.1016/j.hrmr.2015.01.002

Stone-Romero, E. F., Stone, D. L., & Hyatt, D. (2003). Personnel selection procedures and invasion of privacy. *Journal of Social Issues*, *59*(2), 343–368. https://doi.org/10.1111/1540-4560.00068

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In: Worchel, S. and Austin, W.G., Eds.,. In *Psychology of Intergroup Relation* (pp. 7–24). Hall Publishers.

Tews, M. J., Stafford, K., & Zhu, J. (2009). Beauty revisited: The impact of attractiveness, ability, and personality in the assessment of employment suitability. *International Journal of Selection and Assessment*, *17*(1), 92–100. https://doi.org/10.1111/j.1468-2389.2009.00454.x

Tiano, S. (1986). Authoritarianism and political culture in Argentina and Chile in the Mid-1960's. *Latin American Research Review*, *21*(1), 73–98. https://doi.org/10.1017/S0023879100021877

Timming, A. R. (2017). The effect of foreign accent on employability: A study of the aural dimensions of aesthetic labour in customer-facing and non-customer-facing jobs. *Work, Employment & Society*, *31*(3), 409–428. https://doi.org/10.1177/0950017016630260

Torres, E., & Gregory, A. (2018). Hiring manager's evaluations of asynchronous video interviews: The role of candidate competencies, aesthetics, and resume placement. *Faculty Scholarship and Creative Works*, *75*. https://doi.org/10.1016/j.ijhm.2018.03.011

Tulchin, J. S. (2010). The strategic culture of Chile. *Revista Del CESLA. International Latin American Studies Review*, *2*(13), 433–438.

Turner, J. C., & Oakes, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, *25*(3), 237–252. https://doi.org/10.1111/j.2044-8309.1986.tb00732.x

Umphress, E. E., Simmons, A. L., Boswell, W. R., & Triana, M. del C. (2008). Managing discrimination in selection: The influence of directives from an authority and social dominance orientation. *The Journal of Applied Psychology*, *93*(5), 982–993. https://doi.org/10.1037/0021-9010.93.5.982

Undurraga, T. (2015). Neoliberalism in Argentina and Chile: Common antecedents, divergent paths. *Revista de Sociologia e Política*, *23*, 11–34. https://doi.org/10.1590/1678-987315235502

Veit, S., & Thijsen, L. (2021). Almost identical but still treated differently: Hiring discrimination against foreign-born and domestic-born minorities. *Journal of Ethnic and Migration Studies*, *47*(6), 1285–1304. https://doi.org/10.1080/1369183X.2019.1622825

Walker, H. J., Helmuth, C. A., Feild, H. S., & Bauer, T. N. (2015). Watch what you say: Job applicants' justice perceptions from initial organizational correspondence. *Human Resource Management*, *54*(6), 999–1011. https://doi-org.library.smu.ca/10.1002/hrm.21655

Wexley, K. N., & Nemeroff, W. F. (1973). The effects of racial prejudice, race of applicant, and biographical similarity on interviewer evaluations of job applicants. *Proceedings of the Annual Convention of the American Psychological Association*, 593–594.

Wittlinger, R. (2004). Perceptions of Germany and the Germans in Post-war Britain. *Journal of Multilingual and Multicultural Development*, *25*(5–6), 453–465. https://doi.org/10.1080/01434630408668918

Waughray, A. (2009), Caste discrimination: A twenty-first century challenge for UK discrimination law?. The Modern Law Review, 72: 182-219. https://doi.org/10.1111/j.1468-2230.2009.00740.x

Zheng, C., & Menzies, J. (2015). Developing global mindsets through the China study programme. In V. Taras & M. A. Gonzalez-Perez (Eds.), *The Palgrave Handbook of Experiential Learning in International Business* (pp. 550–568). Palgrave Macmillan UK. https://doi.org/10.1057/9781137467720_32

Zschirnt, E., & Ruedin, D. (2016). Ethnic discrimination in hiring decisions: A meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies*, *42*(7), 1115–1134. https://doi.org/10.1080/1369183X.2015.1133279